ARPA ORDER NO.: 3478
9D20 Cybernetics Technology

R-2382-ARPA
February 1979

Design Considerations for Computer-Based Interactive Map Display Systems

Robert H. Anderson, Norman Z. Shapiro

A Report prepared for

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

Rand
SANTA MONICA, CA. 90406

ARPA ORDER NO.: 3478
9D20 Cybernetics Technology

R-2382-ARPA
February 1979

Design Considerations for Computer-Based Interactive Map Display Systems

Robert H. Anderson, Norman Z. Shapiro

A Report prepared for

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

Rand
SANTA MONICA, CA. 90406

PREFACE

This is the final report of a one-year project sponsored by the Cybernetics Technology Office of the Defense Advanced Research Projects Agency. The goal of the project was to determine, through thought and informal experimentation, those key features of an interactive map display system (IMDS) that affect both the usefulness of the system and the design and architecture of its hardware and software. The focus has been on potential uses of map display in command and control systems.

The increasing amount of computing power that can be incorporated in graphic display systems allows new modes of interaction between man and machine—for example, modes in which the machine "understands" the data it is displaying and can take actions based on the content of those data. The techniques and design principles discussed in this report are applicable to command and control systems ranging from support of a field commander through systems tailored to the needs of the National Command Authority.

Although it contains some general discussions of maps and their uses, the report is written primarily as a set of guidelines for designers of computer-based interactive map display systems. It should be of greatest interest to persons concerned with the specification, design, or acquisition of such systems.

SUMMARY

Maps play a fundamental role in planning and decisionmaking activities related to command and control. The abstractions used and the cartographic decisions that result in traditional paper maps have evolved over thousands of years and provide a compact and highly useful representation of geographic information.

The essence of a map is abstraction: Maps generally present a highly abstract representation of reality. They are abstract in that information is omitted (for example, minor roads may not be shown), they are stylized (a city may be represented by a single dot), and they are encoded (the population of a city may be denoted by the size and shape of the dot).

Traditional paper maps must contain all the information that a user is likely to need; hence there is a continual need for cartographers to balance the amount of clutter with a user's need for information.

Although paper maps have many advantages—for example, they are inexpensive and highly portable—they have many disadvantages as well: They cannot represent rapidly changing information; they omit needed information; they do not incorporate useful computational aids such as minimum-path algorithms or time-of-flight calculations.

Recent developments in computer graphics and the continuing decline in the cost of electronics offer the possibility of creating interactive map display systems (IMDSs) having many of the advantages of traditional maps, but significant additional advantages as well.

These systems are not merely maps but are aids to geographic planning and problem-solving in the broadest sense. Experimentation on such systems conducted by the authors during the past year has produced the following observations and guidelines for developers of IMDSs:

- Due to their fundamental differences, the design of electronic maps should not mimic that of paper maps. Each cartographic decision or design feature must be reconsidered based on its underlying purpose.
- In a system with reasonable computational agents, continuous display controls such as knobs or joysticks are in many situations considerably inferior to discrete controls such as function buttons.
- Users have preconceptions that are quite uniform about the direction in which continuous controls should be moved to shift the display. In some cases, the expected direction changes when the size of the object being viewed exceeds the size of the display window.
- Aircraft-type controls are inappropriate for almost all geographic display applications.
- Aircraft-type controls can be learned in a few minutes by most subjects.
- Given control over clutter, users act responsibly and limit clutter effectively.
- Given a choice, users often prefer to receive voice-output data rather than CRT-presented data.

- Disorientation can be caused by abrupt changes in view, lack of visible features, and interruption of the user.
- Disorientation can be reduced by specific training of users and occurs less frequently among trained pilots than among other subjects.
- A variety of simple techniques can be used to retain user orientation in map displays; disorientation is therefore not a significant problem.
- Discrete zoom or translation increments greater than certain limited values cause disorientation and should be avoided.
- Continuously displayed legends giving names of entities (and displayed text giving other properties of geographic entities) seem less valuable in interactive maps than in paper maps.
- Electronic map index programs can greatly increase a user's ability to locate information. They depend, however, on a system design in which the computer does not act just as a camera but understands the names and attributes of the data being displayed.
- Users can tolerate and effectively use variable abstractions, provided they can control the abstraction process.
- Users' abilities to tolerate variable abstractions increase with experience.
- Interactive maps are effective problem-solving devices, even when supplied with rudimentary computational and information retrieval facilities.

These guidelines are not statistically valid experimental results. They result from informal experimentation, with a variety of software and hardware systems. Some guidelines confirm, and some disconfirm, various of the authors' preconceptions. Many guidelines concern the subjective "feel" of an interactive map display system and could not be derived by rational analysis alone.

The results of the experimentation conducted to date are sufficiently promising that interactive map display systems should be considered a viable candidate for many command and control applications. Careful attention to design details and to the true nature of maps as problem-solving aids can result in systems that significantly increase user effectiveness for a variety of planning and problem-solving tasks.

ACKNOWLEDGMENTS

Robert Gammill and Phyllis Kantar of the Rand staff provided significant system programming support for the experimentation upon which this report is based. Mickey Conte, Glen Fleck, R. Stockton Gaines, Barbara Hayes-Roth, Gary Martins, and Doris McClure all provided assistance and advice.

We are also indebted to Frank Lewandowski of Singer-Link Division, Sunnyvale, for his time and for providing access to Singer-Link's sophisticated interactive graphics display equipment; and to Tom Ferrin of the State University of California at San Francisco, Department of Pharmacology, for allowing us to use his graphics software package.

The portion of the map of metropolitan Los Angeles reproduced in Color Plate 2 is used with the permission of the Automobile Club of Southern California.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

Section

I. WHAT IS A MAP? 1
   What Is a Map For? 2
   Some Basic Assumptions 4
   Three Interactive Map Display Systems 5
   Overview of Methodology Used in this Study 9

II. DESIGN CONSIDERATIONS REGARDING INTERACTIVE MAP DISPLAY SYSTEMS 11
   General System Architecture 11
   Five Dimensions for Map Display System Options 12
   Summary of Design Considerations for Interactive Map Display Systems 19
   Implications for System Architecture 33

III. CONCLUDING REMARKS 38

Appendix: MAJOR INTERACTIVE MAP DISPLAY ENVIRONMENTS USED IN RESEARCH EXPERIMENTATION 39

REFERENCES 45

I. WHAT IS A MAP?

Maps and map-like materials remain the basic tools used for geographic problem-solving in command and control. The map is one of the truly remarkable human inventions. It is a profound combination of the abstract and the concrete; it functions as an information storage and retrieval device, as a short-term memory, and as both a digital and an analog computer. It is also a work of art. The modern military map puts into the hands of a commander the synthesis of results from our most advanced and exotic technologies—space, optical, computer, and graphic production; the focusing of vast organizational efforts; and the results of intelligence analysis.

The map materials available for command and control purposes do not differ fundamentally from maps dating back 4000 years. Yet many of the information analysis and dissemination requirements of modern command and control situations cannot be met by traditional printed paper maps. It is therefore important to understand the extent to which computer-generated displays combining symbolic and pictorial information under user control can meet such requirements. But before we transfer maps onto electronic media, we must first ask the question, What are the essential qualities of a map?

We use the following definition: *A map is a partly abstract, partly concrete, graphical representation of geographic data.*

Consider the satellite photograph of the Los Angeles basin shown in Color Plate 1. It is a beautiful object and a tribute to technology, but according to our definition it is an extreme form of a map, since its content is almost totally concrete. Color Plate 2 shows an Automobile Club of Southern California road map of the same area. For its intended purpose, the road map is much more useful because it abstracts much of the concrete reality captured by the photograph.

Although the map reader seldom thinks about the extent to which reality has been abstracted in the maps he uses, abstraction is central to maps. For example, note the drastically simplified road network in the Automobile Club road map, compared with the satellite photo. Abstraction also includes a process of classification; for example, different categories of roads are shown by different widths, colors, or symbols. In addition, abstraction includes the use of caricatures or exaggerations, such as the exaggerated freeway interchange ramps in the road map.

Our definition of "map" involves the concepts of abstraction, concrete representation, and geographic data. The interrelationships among these concepts and others closely related to them are discussed below.

Geographic data comprise a set of data items, each having location (plus possible spatial distribution) as one of its important attributes. Such items include, but are not limited to, terrain, politics, weather, climate, demography, flora and fauna, geology, culture, the disposition of military forces, and rates of change of all of the above.

Abstraction is a general term for the processes by which geographic data are structured to fit human needs. In the geographic context, abstraction is another name for modeling. (The term is by no means universal; cartographers often refer

to this process as "generalization.") It includes the somewhat overlapping concepts of simplification, symbolization, classification, and induction.

Simplification is the selection of attributes of the data, possibly involving the modification of geographic attributes. For example, to simplify a road network, certain roads are eliminated, and in addition, others are straightened.

Symbolization is the summarizing and coding, by graphical or other means, of certain data attributes, including in particular their comparative significance and relative position. Symbols include text, ideographs, pictographs, caricatures, and cartoons. A caricature is a symbol that distorts some attribute of the underlying reality. A cartoon, for our purposes, is a symbol that changes in time as a means of representing information.

Classification is the ordering or scaling and grouping of geographic phenomena. It is this grouping which facilitates human assimilation of the full complexity of reality. The resulting classification of data may of course be represented by the use of symbols.

We have given considerable emphasis to the process of abstraction as being central to the creation of maps. Much of the content of this report deals with various forms of abstraction that we have found useful in interactive map displays. We dwell on abstraction to this extent partly because we believe that the abstractions used to represent information (e.g., symbols) affect our perception of and ability to manipulate that information. In thinking about a DNA molecule, for example, a reader might well visualize a diagram out of *Scientific American*. Needless to say, that representation, although very useful, is highly abstract and far from the underlying concrete reality.

We find the classical literature on cartography to be only of general interest in considering the role of abstraction in interactive map displays; much of this literature is tied too closely to the medium of printed maps. Good introductions to classical cartography for map display designers are presented in Lawrence [1] and Raisz [2]. A recent book on maps which we found interesting, stimulating, and thoughtful (although we do not necessarily endorse all the concepts therein) is Robinson and Petchenik [3]. Topfer [4] is said to be the best available text on classical abstraction. We also found many of the discussions in Davis and McCullagh [5] (particularly the Bickmore article) worth reading.

WHAT IS A MAP FOR?

We are not able to offer a definitive taxonomy of the uses of maps, even for the limited area of command and control. Several examples illustrate the diverse activities in which paper maps traditionally play a role:

- *Geographic problem-solving:* Finding the minimum-time route for a battalion in transit from Calais to Charleville while avoiding enemy concentrations at Lille; drawing three radio direction-finding azimuths to see where they intersect and estimating the circular error probability from their lack of mutual intersection; looking at the ground coverage swaths of various sensor systems and planning routes for flyovers of unallocated sensors to completely cover a geographic area.

- *Information retrieval:* "Where are all the radar sites in sector 12?," "What's the name of this hill?," "Where's the hill named B243?"
- *Information storage:* Drawing in the location of new radar sites as the information is received.
- *Analog computation:* Measuring the angle between a drawn azimuth to the enemy artillery location and the azimuth to a known location.
- *Digital computation:* Counting the total enemy troop strength and firepower shown in the order of battle information displayed for a sector; adding up the kilometers for each segment of a planned route of march to determine total route length and traversal time.
- *Direct person-to-person communication:* "I'll reach *this* bridge by 1330, then take *this* unimproved road to *that* intersection . . ." (with the speaker pointing at map locations *here, there,* and *there).*
- *Indirect communication:* Marking your route of advance, and sending the map by courier to headquarters.
- *Personal orientation:* If that water tower is *this* one, then I should be *here,* and that direction is north . . . .
- *Group orientation:* "You will enter the city *here,* then proceed directly down Burgstrasse to the ammunition dump *here;* watch for snipers on top of *this* building."
- *Gestalt perceptions:* "It looks as if they're massing strength for a major assault." "Judging by the areas colored in red, I'd estimate the enemy is controlling about 30 percent of district 7."
- *Memory aid:* Drawing your patrol's route on a map and sticking it in your pocket.
- *Presentation of information for persuasion:* During the Yom Kippur War between Israel and Egypt, it is reported in Safran [6] that President Sadat of Egypt was not convinced of the need to pursue a cease fire until shown Soviet satellite photographs, whose summary in textual form he had previously seen.

It is probably more interesting to consider some activities that traditional paper maps are *not* good for:

- *Asking questions about the displayed data whose answers were not explicitly encoded in the map symbolism:* "Where are all the bridges in sector 13 capable of supporting a 12-ton tank?"
- *Providing an up-to-date display of a rapidly changing situation:* This is often accomplished in a manpower-intensive fashion with greaseboards and personnel skilled at writing backwards with grease pencils.
- *Geographic problem-solving involving substantive computations:* "What time should the aerial refueling take place *here* if our bomber squadron follows this route . . . ?"; displaying the computed ground coverage swaths of all in-place sensor systems, then drawing in reconnaissance routes with their associated swaths to provide complete coverage of a geographic area.
- *Showing the imprecision or staleness of information:* "Who drew in that SAM site? Peterson? Where did he get that information?"

- *Showing time-varying information, such as the coordinated pattern of enemy troop movements over the last 48 hours.*
- *Concise, accurate communication among geographically dispersed parties:* A: "I'll take out *this* target and *this* one, if you can provide covering fire *here.*" B: "OK, but approach from *this* direction to mask your intentions and provide me a clear field of fire." C: "I approve of that plan; let's do it."
- *Allowing different people with different tasks to communicate:* The G2 (intelligence) map display shows restricted data, but certain data elements are received from, and sent to, the G3 (operations) display. As those data elements are updated by the relevant staff component, they are updated simultaneously on all other staff displays as appropriate.

We have emphasized two activities in creating the above list: geographic planning and problem-solving (and planning is really just a form of problem-solving). Since the design of an interactive map display is dominated by its intended uses, this report is really about geographic problem-solving aids, not maps. We shall therefore emphasize aids to geographic problem-solving that capitalize on the interplay between the computational power, possible communication links, information storage, and the display.

SOME BASIC ASSUMPTIONS

Our study of computer-based geographic problem-solving systems is based on a number of assumptions. Taken together, these assumptions lead to the belief that within the next five to ten years it will be possible to construct cost-effective interactive map display systems; these systems will be able to handle many traditional map functions (the first list, above) and will also provide significant help in the areas (in the second list) poorly supported by traditional paper maps. Some of our major assumptions are listed below, in no particular order:

- Maps are a ubiquitous, important means of storing geographic data for use in familiarization, planning, and problem-solving activities in command and control.
- Computer display hardware will soon have the flexibility and power required to display map data in a way different from, but more effective than, traditional static paper maps; these displays will also have the capacity to be updated and changed rapidly so that truly interactive man-machine systems can be developed.
- Displayed maps are very different from static paper maps. For example, (1) an interactive system must to some extent act as a cartographer; (2) cartographers designing paper maps must balance clutter with the omission of needed information, whereas an interactive system can selectively display from a potentially vast data base of information, all of which need not—or cannot—be viewed at one time; (3) interactive maps can use three-dimensional and dynamic symbols; (4) the interactive map can be used as an interface to other data-processing services, or as a communication tool.

From these assumptions, we derived the following research hypotheses that guided our study:

- Displayed maps are potentially much more valuable and powerful as a problem-solving and planning tool than paper maps. They can show updated and changing information in a dynamic situation; their displays can show the results of complex computations; their symbology can indicate the timeliness and certainty of information, thereby creating a more accurate impression for the user of the data available for decisionmaking.
- Since interactive map display systems are so different from traditional maps, their attributes should be explored through informal experimentation prior to the development and field-testing of major prototype systems.

THREE INTERACTIVE MAP DISPLAY SYSTEMS

What would a complete interactive map display system look like? We present below brief sketches of three representative interactive map display systems, each in a different context, and discuss the design choices we feel are appropriate for each context. Obviously, any particular system developed for installation and test in a particular operational environment (even in these context areas) might well have unique requirements not discussed here that bias the system design. Our remarks should not be taken as specific design recommendations for IMDSs for these contexts, but merely as "gedanken experiments" in the choice of display system parameters to meet certain generic requirements.

Since these system sketches are introductory in nature, some of the vocabulary may be unfamiliar to some readers. The concepts and system attributes mentioned here are discussed in greater detail in Sec. II.

The three representative systems considered are: air interdiction planning, theatre air target selection, and a National Command Authority intelligence situation briefing display.

System 1: Air Interdiction Planning System

A sophisticated display system to aid in decisionmaking and planning involving three-dimensional air operations.

Air interdiction operations involve the choice of geographically dispersed targets behind the enemy front line and the selection or creation of "corridors" to be used as main flight routes to the target area. Corridor selection involves complex three-dimensional information, such as the location of radar "beaten volume" cones to be avoided, air defense perimeters, satellite coverage swaths, etc. We propose a system to be used by an interdiction planner that will allow normal map operations in target selection but that will, in addition, allow him to "fly" a route in three-dimensional space, providing a pilot's-eye view of the obstacles (electronic or otherwise) to be faced on that route.

We propose the following design decisions; they are discussed in terms of five design "dimensions" with which we characterize map display systems. The numbers shown in the center column refer to the numbered paragraphs in our discussion of each design dimension in Sec. II.

	Dimension	Item in Sec. II	Design Choice
1.	Perspective control	(4)	Real-time "flying" of display for perspective and scale control.
2.	Information generation, manipulation, retrieval, and storage	(5)	Multiple information overlays, with possible time-varying information; some automatic clutter control.
3.	Abstraction control	(1)	No user control over abstractions.
4.	Input/output modes	(4)	1024 $\times$ 1024 color CRT with mirror projection optics; two aircraft-type stick controls for dynamic perspective control; possible tablet for menu selection.
5.	Communication	(3)	System receives formatted data base update information over low-bandwidth link to reflect latest intelligence reports on target and radar locations.

Figure 1 shows a sketch of a representative display scene created by this hypothesized system. This system emphasizes the use of three-dimensional perspective to allow planning in three-dimensional space.

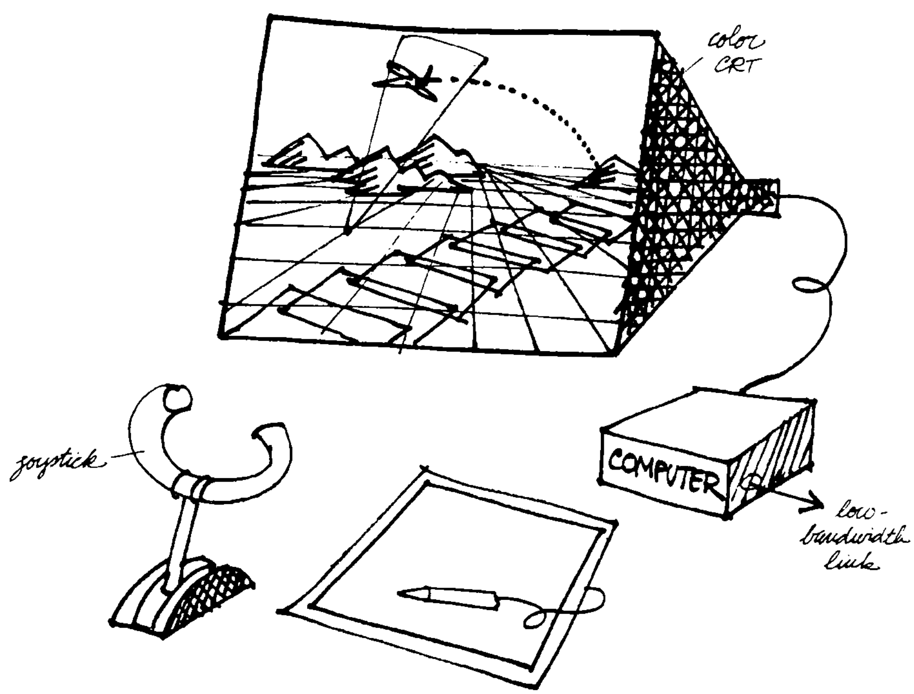


Fig. 1—System 1

System 2: Theatre Air Target Selection Display

A modest-cost system providing some advantages over traditional paper maps for target selection purposes.

This application is a subset of System 1; in this case, we are concerned only with the selection of targets for air interdiction behind the enemy's front line. We are not concerned with corridor selection involving three-dimensional route planning. We assume that a "bird's-eye" view of a rather traditional map display is sufficient and that different map scales can be chosen in discrete increments (such as plus or minus a factor of 2). Similarly, translation (i.e., sideways movement) of the map display is in discrete increments. This is a stand-alone system and is not used for communication.

We propose the following design choices:

	Dimension	Item in Sec. II	Design Choice
1.	Perspective control	(1)	Bird's-eye view only, with discrete zoom levels and discrete translation increments.
2.	Information generation, manipulation, retrieval, and storage	(3)	Fixed library of named, stackable overlays; combinations can be displayed or stored for later retrieval.
3.	Abstraction control	(1)	No user control over abstractions.
4.	Input/output modes	(1)	Standard 512 × 512 color CRT with keyboard.
5.	Communication	(1)	No communication with external systems; periodic update of data base by replaceable cartridge or disk.

Figure 2 shows a representative display from this hypothesized system. The display is quite similar to a traditional map, except that a potentially much larger data base can be accessed, and there is more freedom in building and manipulating overlays.

System 3: Intelligence Situation Briefing Display for National Command Authority

A flexible, powerful display system that relies on a trained specialist—an intelligence briefing officer—to manipulate the display options for maximum effectiveness.

Intelligence briefings at the National Command Authority level—including briefing of the President—are performed by briefing specialists. Time is limited, and considerable information must be imparted succinctly and accurately, through words and graphics. A good model of the current situation must be created in the minds of the audience. Cost is not necessarily a constraint, if effectiveness is demon-

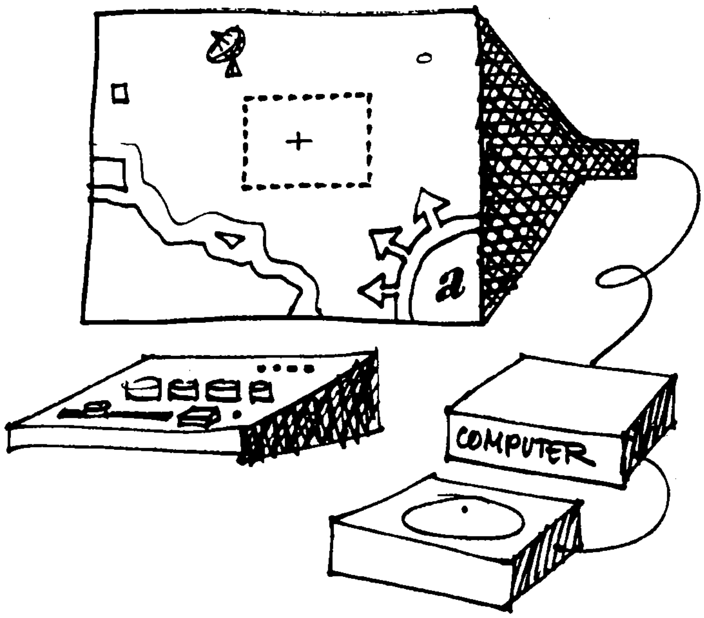


Fig. 2— System 2

strated. The set of design choices shown below provides a system with considerable flexibility; we assume that the briefing specialist can manipulate the display system features to provide a rich information environment.

	Dimension	Item in Sec. II	Design Choice
1.	Perspective control	(2)	Bird's-eye view only, but with the possibility of continuous control over the magnification (scale) and translation of the map within larger discrete increments; use of electronic pointer to focus attention.
2.	Information generation, manipulation, retrieval, and storage	(10)	Complete information retrieval language; use of parameterized information overlays to show time-varying information.
3.	Abstraction control	(6)	Intelligence officer and his staff prepare appropriate symbology (not necessarily a simple task; could involve some programming).
4.	Input/output modes	(2)	High-resolution CRT, keyboard, some simple knob-type controls for limited zoom and scroll.
5.	Communication	(2)	Vast bulk of information prestored; some limited communication possible for dynamic update of information from remote sources.

Figure 3 shows a representative display from this hypothesized system. The display uses more elaborate abstraction and symbology to condense and represent a large amount of information.

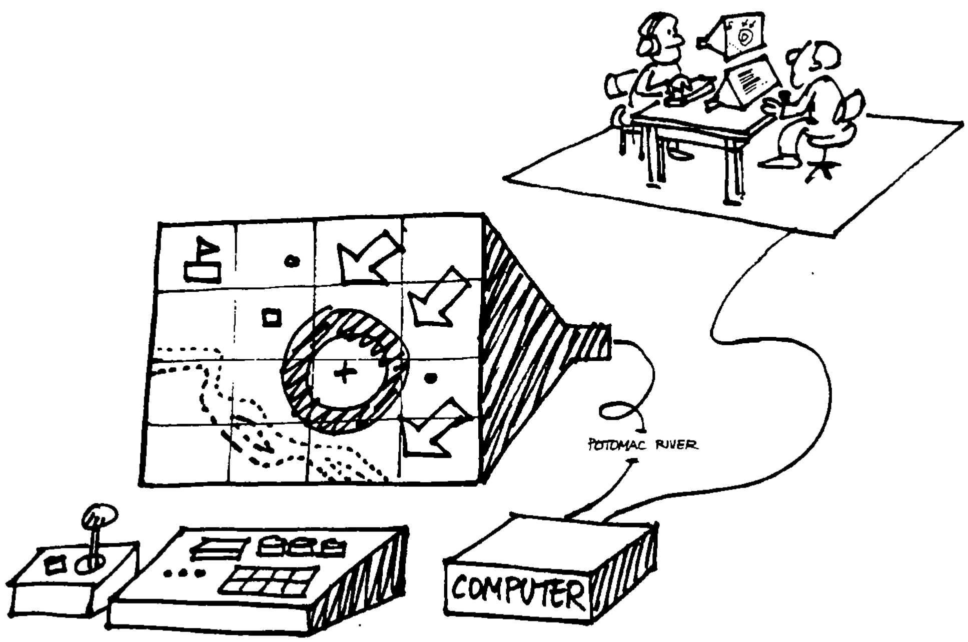


Fig. 3— System 3

OVERVIEW OF METHODOLOGY USED IN THIS STUDY

We believe the essential quality of a map is the form of abstraction and symbolism used to represent information. Consequently, we have concentrated our resources on understanding the new options for map symbolism opened up by interactive graphics.

We conducted several *informal* experiments (described in the Appendix) to test the effectiveness of various forms of symbolism in problem-solving situations. We stress the word "informal." Although a number of persons not directly involved with the project were exposed to our interactive software and were asked to perform various tasks using it, these investigations were exploratory and were not controlled or statistically valid. We looked for gross effects—factors of 2 or 4 or 10 in increased effectiveness.

This report does not describe experimental procedures or experimental results, nor does it examine particular hardware or software systems. Moreover, it does not review currently available map display systems, since their design is understandably based on existing hardware and software limitations and presents too constrained a view of the future. Rather, we present a set of design guidelines regard-

ing interactive map display systems. These guidelines cannot be proven as design principles, but each is strongly felt by the authors to have general validity. Each is based on personal experience—and observations of others' experience—with interactive software. Some of the guidelines contradict, and others confirm, some of our initial preconceptions.

We should mention here one other subject this report does not address: computer-generated realistic views of objects. With current technology, it is possible to perform the necessary hidden-surface computations and shading algorithms to allow a user real-time control over views of a realistic-looking scene. (We have in fact experimented with such a system at Singer-Link Division, Sunnyvale.) These displays are highly appropriate for flight training simulators and related activities. But although these realistic displays might be considered to be "movie maps" and although they have relevance for such activities as briefings regarding terrorism incidents (assuming that the specialized data base required can be created or will become available), we do not believe they are relevant for the normal spectrum of activities for which maps are used; they are neither sufficiently selective nor encoded to make information traditionally found in maps available to the user. We therefore do not address this specialized use of computer graphics.

It is difficult to form objective conclusions about the design of interactive man-machine interfaces. Some of our guidelines (for example, our preference for discrete and content-based zoom control over continuous control knobs) are based on the feel of the system, and the fact that one way may take 1 second to do something that takes 8 seconds another way. The importance of the difference of 7 seconds depends on the context in which the operation occurs and the frequency of its occurrence. We despair of providing objective proof; if a guideline seems irrelevant or wrong, we can only say, "If we have made a videotape or film by the time this report is available, ask to see it; better yet, you must experiment with a diverse set of relevant interactive software yourself."

II. DESIGN CONSIDERATIONS REGARDING INTERACTIVE MAP DISPLAY SYSTEMS

GENERAL SYSTEM ARCHITECTURE

In this report, we consider an interactive map display system (IMDS) to be a relatively self-contained unit comprising at least (1) a local data base containing geographic and other command and control data; (2) a display system (such as a cathode-ray tube (CRT)) and associated logic; and (3) computational power provided by one or more computers capable of both the arithmetic computations needed for display handling and data base storage, retrieval, and updating operations. In addition, the system must interpret a variety of user input devices, such as a data tablet, keyboard, or joystick. The IMDS will most likely have one or more communication channels to external systems, either to other IMDSs or other information systems within a distributed command and control network.

The general architecture of an IMDS is represented in Fig. 4. There are a number of questions that might be raised regarding this architecture, such as:

- How much computational capability is needed, and at what cost and physical size can it be provided?
- How large a local data base is necessary for representative applications? How large a bandwidth is needed between the data base and the computer controlling the display?

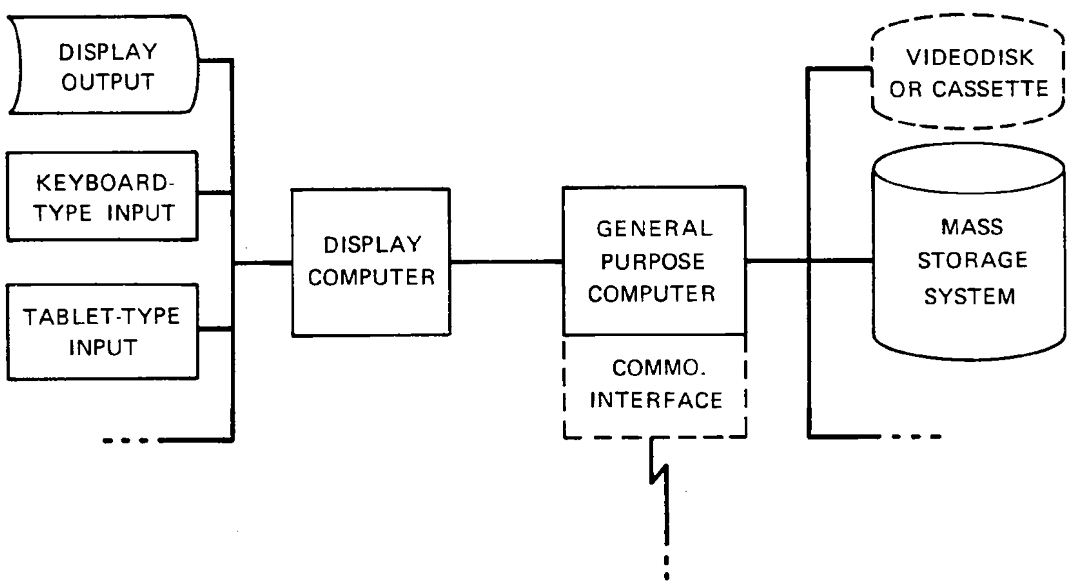


Fig. 4—General architecture of an interactive map display system

- How should the large geographic data base required for map display be organized to allow efficient access?
- What user input modes are most likely to be used, and what computational requirements do they impose?
- What are the forms of communication with external information systems? How large a bandwidth is needed? Is it a continuous or sporadic requirement?

We have not addressed such architectural considerations directly, since our investigation is primarily concerned with the display of geographic abstractions under the control of a user performing problem-solving or planning activities. Nevertheless, when our conclusions bear on architectural issues, we have attempted to describe those relationships.

FIVE DIMENSIONS FOR MAP DISPLAY SYSTEM OPTIONS

A design guideline is a choice among options, selected from a wide range of design possibilities. To interpret a guideline (such as, "Use a data tablet for x"), the reader must know the options from which a selection has been made (for example, basic user input devices include keyboards, data tablets, continuous control knobs or joysticks, function buttons, and light pens). This section provides a general overview of what we have chosen to view as a five-dimensional space of design choices regarding IMDS characteristics. By first outlining this space of options, we hope to provide a context for the discussion of design guidelines that follows.

We find it useful to characterize the options available in the design of an IMDS along five different dimensions. These dimensions certainly do not exhaust the design possibilities, but they capture many of the important options. Each dimension is viewed as a rather continuous spectrum; we have selected, numbered, and discussed certain points along the spectrum as being of special interest. Although these five dimensions are not strictly "orthogonal"—that is, mutually independent—we nevertheless informally consider them as defining a five-dimensional space of design options, with particular systems occupying a point in that space.

The five dimensions are listed below. Each ranges from little or no user control over the option up through considerable—usually real-time dynamic—control.

1. *Perspective control:* The amount and type of user control over the way he views the displayed data: scrolling, zoom, rotation, and viewpoint selection.
2. *Information storage, retrieval, and manipulation:* User control over the retrieval of information from the IMDS data base, its manipulation and display on the electronic map, and the ability to store newly created map displays for later retrieval.
3. *Abstraction control:* User control over the symbols used to represent information, and such attributes of those symbols as their time-varying nature.
4. *Input and output modes:* The degree of flexibility and power in the input and output devices and modes available to a user.

5. *Communication with external systems:* The degree of communication capability of the IMDS, ranging from none through extensive real-time exchange of information with external information systems.

Each of these dimensions is discussed in more detail below. The numbers attached to points of discussion along these dimensions are for reference only and are not meant as a numerical scale representing absolute position within that dimension.

Dimension #1: Perspective Control

One of the obvious advantages of electronic maps is that the user can change scale at the touch of a button or knob. He can scroll the map to focus the full display screen on a portion of the information. For some applications, the user might want a three-dimensional perspective view of the map, especially if the map itself contains three-dimensional representations of topography, radar or weapon coverage cones, etc. User control over scale change involves a problem not encountered in paper maps: Different abstractions are appropriate for representing the same information at different scales. If a user can continuously and dynamically change the map scale from 1:50,000 to 1:500,000, how can the symbols used to represent marshland, roads, etc. be gracefully changed (or added and deleted) so that they remain appropriate over that entire range? We return to this problem later but introduce the topic at this point to alert the reader to some of the implications of system design choices involving perspective control. Some of the possible degrees of sophistication in providing these facilities are:

1. *"Birds-eye view" only, fixed aspect ratio, discrete zoom levels, discrete translation increments, modest maximum scale.* In this limited option, the user might, for example, be able to change scale only by a fixed increment, such as doubling or halving it. He might be provided two buttons labeled something like "x 2" and "x 1/2." The total range over which scale could be changed would be limited, perhaps to about five factors of 2, giving a total range from the minimum scale to 32 times that scale. Buttons might also be used to scroll the map to adjacent areas such as the eight nearest neighbor "pages."

   There are very sound reasons for choosing such a limited range of user options. Perhaps the primary reason is simplicity of design of the underlying data base supporting the map display system. This option allows total precomputation of the display at each of the fixed scale and translation increments, including appropriate choice of the abstractions used to represent information appropriate to each scale. These fixed displays can be placed on a mass storage device in such a way that access to the relevant block of data is rapid and efficient. As the user is given more flexibility, the requirement for dynamic computation of the display from its underlying data base increases correspondingly, as does the requirement for more general but less efficient storage schemes.
2. *Same as option 1 above, but with some continuous user control over magnification and translation within the discrete levels and extremely simple abstraction changes.* This option retains many of the advantages of the

previous one but gives the user greater flexibility at the expense of electronics or software for dynamic zoom and translation computation. Some continuous control devices, such as knobs or joysticks, are probably required. This option continues to allow total precomputation of the display, including prechosen abstraction.

3. *Continuous zoom control, translation, etc., with some real-time limitations (gracefully applied).* This option provides the user with the ability to continuously change scale or translation. Since it is not possible to prestore all the map display output that can be generated in this manner, the computer must generate abstractions dependent upon such factors as the current scale. Some simplifications (and resultant cost savings) are possible if the display is not required to "keep up" with the user at all times; that is, under some conditions, response to some commands might be delayed to enable computed abstractions to catch up. The delay would probably not be more than one or two seconds of elapsed time.
4. *Same as option 3, but with complete real-time response (e.g., for flight simulation).* This option provides the user with cockpit-type "stick" controls for complete continuous, dynamic, real-time control over all perspective parameters. The user can "fly" through a simulated information environment as if he were in an aircraft. The computer generates all needed abstractions and computations of the appropriate perspective view of these abstractions in real time.

Dimension #2: Control Over Information Generation, Retrieval, Manipulation, and Storage

It is often useful to think of an electronic map as a two-way two-dimensional window onto an information storage and retrieval system. The information system may be queried in either the traditional manner ("Where are all enemy corps headquarters in sector 12?") or through use of the map itself ("What is the name of *this* river?"). Similarly, the responses may be in traditional text form or may take the form of displayed information on the map (e.g., the locations of all enemy corps headquarters might be represented by the appropriate military map symbol, blinking on and off to highlight their locations on the display). There is a range of options regarding the sophistication of the information storage and retrieval allowed and the modes of user access to these facilities:

1. *Choose one of k prestored map systems (where k is small).* In this most limited option, the computer-based system acts as a file and retrieval system for a small set of prestored maps showing useful information. The user can choose to display one of several (e.g., up to 20) prestored "map systems," each designed to be optimal for some aspect of the user's problem-solving activity. (Such map systems might display transportation networks, including roads, rail, and bridges; or dispositions of friendly and enemy troops, including the FEBA and designated areas of friendly and enemy control.) This option allows cartographic decisions to be premade, either manually or with more elaborate off-line computation.
2. *Same as option 1, but with k large.* This option involves a much more substantial data base management system and storage capacity. The val-

ue of k might be, say, 1024. This option retains the advantages of possible precomputation of complete maps, so that no real-time cartographic decisionmaking is necessary.

3. *Same as option 2, but with a fixed library of named, stackable overlays.* Any combination of overlays can be simultaneously displayed. The user can create new named overlays as combinations of existing ones. With this option, the user can create $k = 2^n$ maps, where n is the number of overlays, since each overlay can be independently displayed, if desired. Some computer-based cartographic decisionmaking is probably necessary to avoid clutter problems. This option does not preclude the storage of one or more of the overlays as a static bit-map (e.g., with political and continental boundaries) with no computer-readable semantic content. A more elaborate language becomes necessary by which the user can add and delete particular overlays from the display and create new maps from combinations of previous ones.
4. *Overlays with parameters.* This option is similar to the previous one, but it permits each overlay to be associated with one or more parameters (e.g., time). Whenever a particular overlay is displayed, the information contained on it is updated or changed as some function of that parameter. Note that "time" can mean real (clock) time, simulation or model time, elapsed time, etc. Clutter control and other cartographic decisionmaking must now be performed in real time.
5. *Automatic clutter control based on continuous levels of importance.* This option assumes that information would have an associated "importance" attribute, usually a continuous variable (or a value taken from a large, discrete ordered set). Examples include the population of cities, maximum runway length times number of runways for airports, and manning levels of enemy units. Computational heuristics are used to resolve clutter and other cartographic problems by making decisions regarding what to show the user, based on the relative importance of the various data. As the user increases the scale of the map (i.e., zooms in), additional information can be displayed in order of importance, since clutter will be lessened as some previously displayed information disappears from the edges of the screen.
6. *Simple Boolean combinations of data attributes used for a data retrieval system.* In this option, the display screen contains possible overlays plus symbols representing data responses to queries consisting of Boolean functions of data attributes. For example, the user might first display the political boundary and road network overlays, then type in the data retrieval request:

   SHOW roads AND railroads AND [bridges with [load > 10 ton]
   OR [capacity > 3 lane]].

   (Note: We show reserved words used to frame a request in capital letters, and we use square brackets to delimit the scope of certain connecting words, just as parentheses are used in mathematical expressions to group terms together.) This option requires a full-fledged information retrieval system backing up the map display system, but the restricted query format allows easy computer processing of the input request.

7. *Elaborate information retrieval system.* Here, the screen becomes a communication medium between the user and a full-fledged information retrieval system. A good example of the current state of the art of relevant information retrieval systems having a natural language interface is the LADDER system designed by Earl Sacerdoti and associates at SRI International [7]. Examples of the type of user query that might be supported include, "What is the name of this hill?" or "Show all roads in sector 4 capable of transporting tanks between junction A7 and Warsaw in less than 6 hours." In this and the previous option, we assume that the concept of an electronic map overlay remains useful for data handling and that the user can build up named overlays with data resulting from information retrieval requests, for later use or in combination.

Dimension #3: Abstraction Control

Recall that we use the term "abstraction" to represent all the processes and decisions, such as selectivity and choice of symbolism, by which information becomes represented in a map display. In traditional paper maps, all control over abstraction is exercised by a trained cartographer. In interactive electronic maps, it is possible (but not necessarily always desirable) to give the user some control over the abstractions used to represent information. Some options in this regard are:

1. *No user control.* All symbolism and abstractions used to represent information are built-in and not subject to user control.
2. *User control based on a menu of options.* In this case, the user does not have the ability to create new symbolisms, but he can choose among prestored symbols. Other menu options might include choice of color or size of symbols, etc. The obvious advantage of this and the previous option is simplicity of the software required to support the limited amount of user flexibility in tailoring map display parameters.
3. *User creation of point symbols.* The user can create new symbols, but they must be "point" symbols; that is, their placement is governed solely by a single (x,y) coordinate. For the creation of new symbols, the user might be given a data tablet to be used in a "draftsman" mode, in which the entire display surface is used to draw a new symbol. That symbol is then used in greatly reduced size to represent a particular type of data, such as the installation sites of a new type of radar unit. When the user is limited to creation of point-type symbols, the linkage between a symbol's position on the display and the underlying data is straightforward.
4. *User creation of linear symbols.* The user can create new symbols and associate them with linear information (such as roads, rivers, communication lines, boundaries) having path characteristics. Here the user must be provided with a language for describing the relationships between portions of a newly created symbol and attributes of the underlying data (such as the linear route to be represented as a road).
5. *User creation of area symbols.* Same as options 3 and 4 above, but newly created symbols can be associated with (and have their shape dependent upon) area characteristics of data. Examples might be area under friendly control, or marshland.

6. *User creation of parameterized symbols dependent on data attributes.* User-created symbols can be associated with arbitrary data attributes or computed values (e.g., load capacity of a bridge, population of a city, (continuously changing) time of day). Again, more sophistication is needed in the language by which a user can describe the mapping between data attributes and the parameters associated with the symbols representing those data.
7. *Computer cartography.* In this option, the computer contains numerous heuristics allowing it to make cartographic decisions regarding choice of symbolism. This choice might depend on context—e.g., other displayed information competing for the user's attention. This decision process could be either automatic or user-guided through man-machine interaction. Such computer decisionmaking will most likely need to be performed in near real time, in response to user requests for information retrieval (see Dimension #2, options 6 and 7). The provision of this feature is beyond the current state of the art; substantial R&D would be required before it could become available in a practical form.

Dimension #4: Input and Output Modes

This design dimension involves the devices by which input signals are sent to the map display system by the user and output displays are presented. Input and output are really separate dimensions; for example, we might choose an extremely simple and limited input device but an elaborate and sophisticated output device. However, these functions are combined for brevity in the following options.

1. *Output: simple CRT; input: keyboard.* A simple IMDS might consist of a standard CRT (e.g., 512 $\times$ 512 resolution with several colors) and a standard alphanumeric keyboard, possibly with several auxiliary buttons for perspective control (see Dimension #1, option 1).
2. *Output: high-resolution CRT; input: keyboard with additional simple continuous controls.* Similar to option 1, but with 1024 $\times$ 1024 CRT resolution. In addition to the keyboard, this system has one or more knobs or standard joysticks for continuous display control.
3. *Output: high-resolution CRT, plus limited voice response; input: keyboard, with continuous controls and tablet stylus for both pointing and drawing.* This design option adds voice synthesis from a limited vocabulary (e.g., 500 words) to provide a parallel output channel, in addition to the visual medium. The tablet stylus can be used to control cursor movement, pointing and menu selection, and possibly the drawing of routes or new symbolism.
4. *Output: high-resolution color CRT, voice synthesis, hardcopy option; input: keyboard, continuous controls, tablet, two airplane "flight-control sticks" for simultaneous dynamic control of approximately six parameters.*

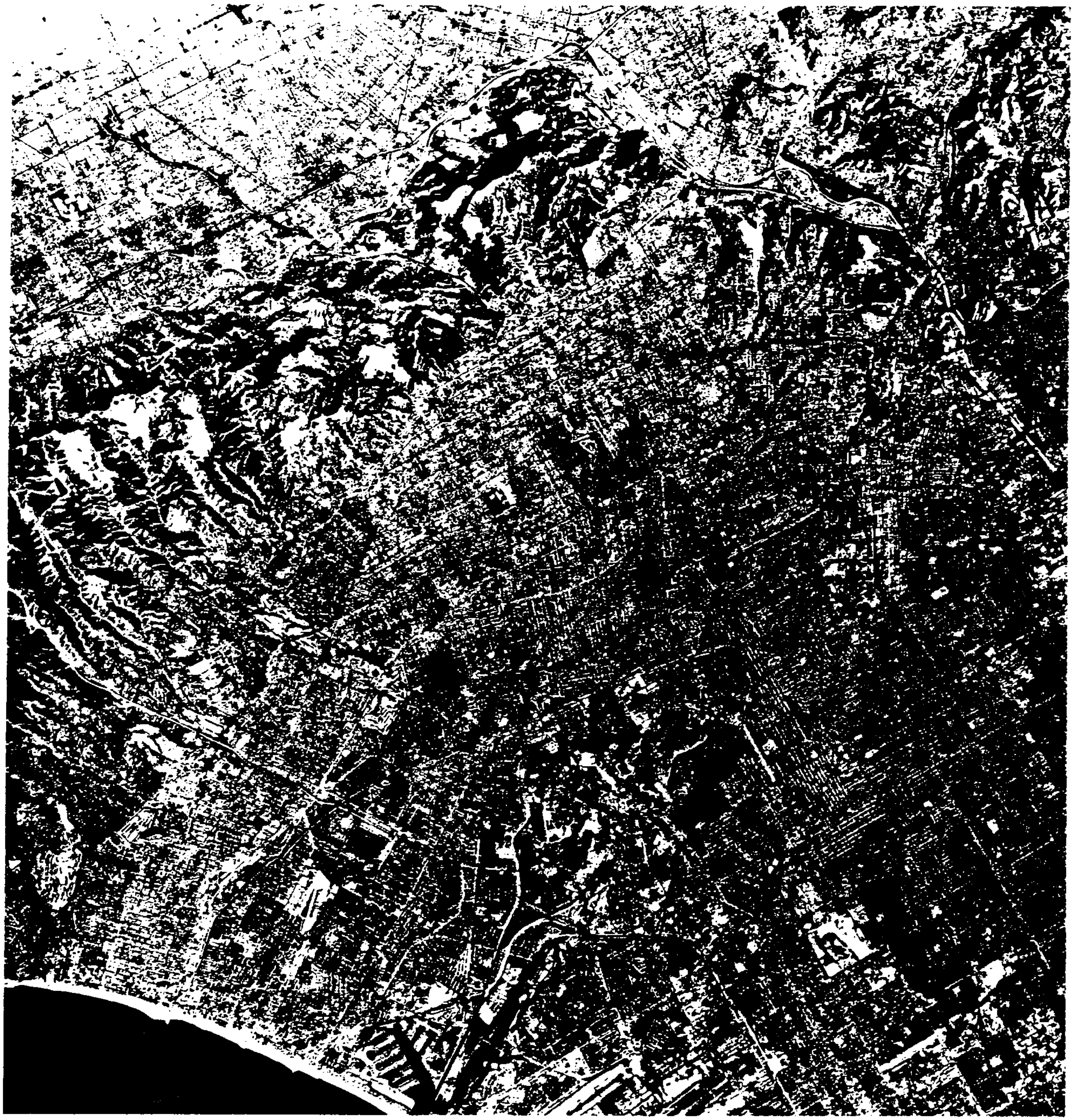
The use of voice synthesis overcomes the vocabulary limitations in option 3 above. Aircraft flight-control sticks allow the user to "fly" through three-dimensional space for perspective and scale control. This design option might well include some tactile input/output, such as programmable resistance or inertia in knobs and

other continuous controls. The hardcopy output option allows local production of paper maps containing the contents of the display screen.

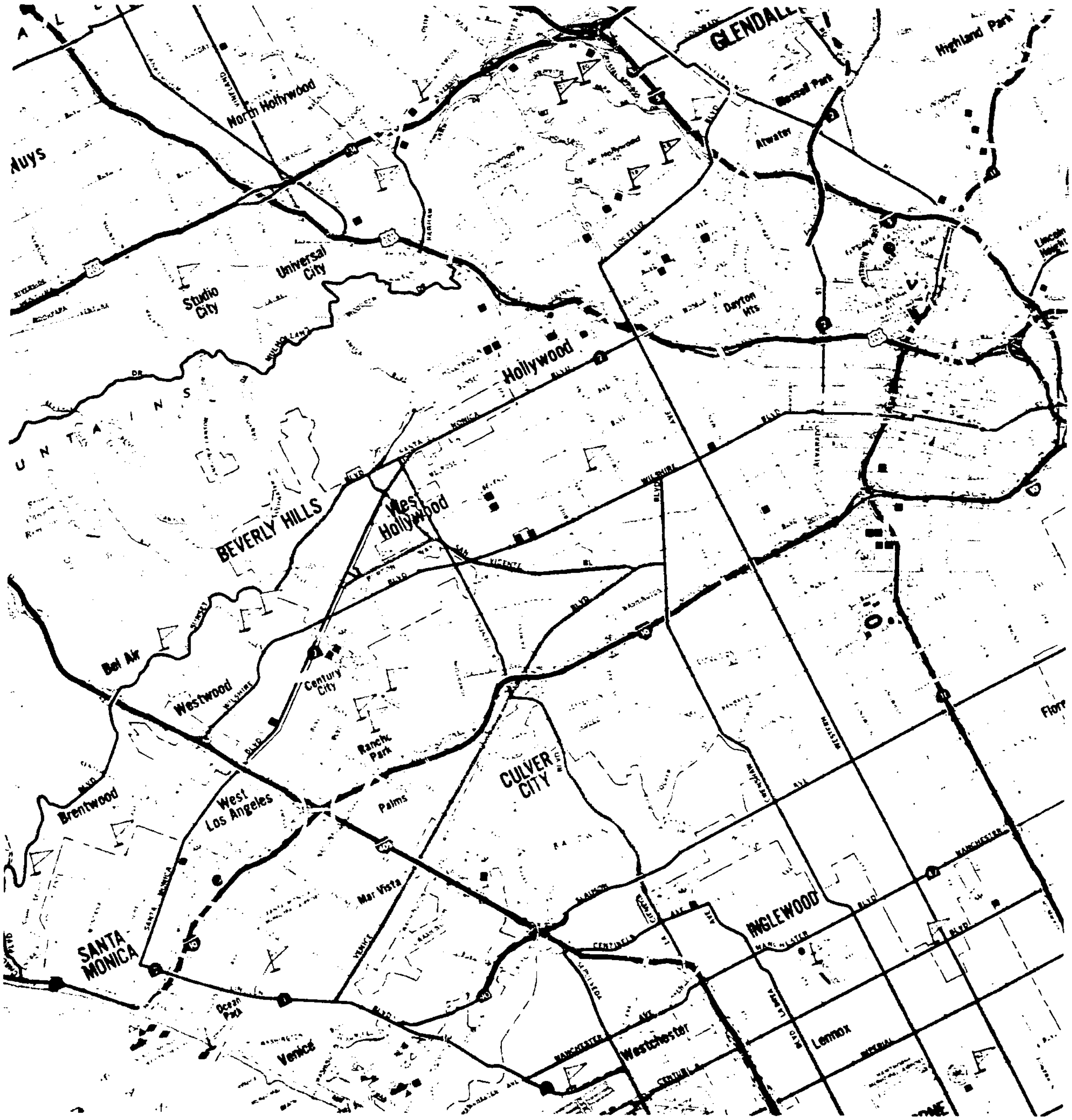
Dimension #5: Communication with External Systems

In many applications (for example, with "intelligent" office automation terminals), a computer system in isolation, removed from communication with other similar systems and external sources of data, is sterile. This same maxim might well hold for interactive map display systems. If so, the amount and type of communication capability in an IMDS is an important design attribute. In some contexts, the IMDS might even best be thought of as a communication terminal, with the system's primary purpose being coordination and communication among geographically dispersed sites. Some of the options are:

1. *No communication.* The map display system is a stand-alone device not capable of communicating with external persons or systems. It is presumably "loaded" with data by a replaceable cartridge or disk pack.
2. *Low-bandwidth communication with other IMDSs using a standard intersystem message sending and receiving protocol.* In this option, the IMDS uses a standard "message-like" protocol to receive and send data base update requests. For example, it can receive a formatted message giving new (x,y) coordinates for a labeled radar site. Similarly, it can send, in message format, a limited stream of (x,y) positions to represent a path drawn by the user (e.g., in response to a request from higher headquarters to show the planned route of advance). (Note that a low-bandwidth communication channel is sufficient for many purposes; a message regarding a route of advance need only contain a sequence of several dozen (x,y) coordinate pairs, allowing the display of a route to be generated from this information at a remote headquarters map display. It is not only unnecessary to transmit the entire map, but the transmitted information can be viewed at a different scale than that in which it was generated.)
3. *Data base update and change commands exchanged with other heterogeneous information systems.* The map display system can generate a richer interactive protocol for exchanging information with other data systems. Through one or more standard information transmission protocols, a number of such systems are capable of maintaining the consistency of several disparate data bases.
4. *Text send and receive capability.* The map display system can act as a general originator and receiver of free-form text messages. Standard text messages might well be generated automatically by the system for external human consumption in predetermined circumstances. (For example, if a forward observer draws the standard symbol for an enemy SAM site on his map display system, his system automatically generates the prestored text message, "From user 0131: enemy SAM site located at 058230/328550 at time 0830Z.") Received messages might be stored, possibly prioritized, and scanned for keywords indicating relevant information for the user's attention or some automatic system display action.
5. *Sensors and controls.* In this option, the map display system is in continuous communication with one or more sensors or control systems. The



1. Infrared satellite photograph of the Los Angeles basin



2. Road map of the Los Angeles basin (same area as shown in Color Plate 1)

3. Computer-generated display of a portion of downtown Las Vegas

system can be used to target sensor systems or to control remote vehicles under user guidance using map display facilities. If aircraft-like input devices are available (see Dimension #4, option 4), a simulated topographic display can be used to "fly" a remote vehicle. Continuous (or sporadic) sensor information might be received and used to update the map display dynamically (e.g., by controlling position, animation, or other parameterized displayed abstractions; see Dimension #2, option 4).

6. *Rapid generation of data bases for crisis applications.* This option entails the preestablishment of technological and organizational mechanisms to rapidly generate data bases from diverse sources, for display of geographic information at very large scales. Crisis management teams might use such a system to quickly generate a tailored map display for a unique situation for which existing map displays were inadequate or inappropriate. Individuals skilled at obtaining information from a variety of disparate data bases would automatically reformat and tailor the information for use in a map display system. Communication links from the IMDS to these data sources would be used to build a tailored geographic data base under severe time constraints. To the authors' knowledge, this capability —both technologically and organizationally—has not yet been demonstrated.

SUMMARY OF DESIGN CONSIDERATIONS FOR INTERACTIVE MAP DISPLAY SYSTEMS

Our five-dimensional design framework contains a wide range of options, many of them not appropriate for many applications. Given this rich space of design options, what choices should be made? This section presents the particular recommendations and guidelines that have resulted from our study to date.

We have reached a number of conclusions (some of them tentative, but all based on first-hand experience and analysis) regarding the design of effective interactive map display systems. We have alluded to some of these conclusions in the previous section. All the guidelines and "results" discussed below are to be interpreted as indications, not as firm doctrine. Much more remains to be learned about the new medium of interactive map display; these guidelines are only a beginning. Nevertheless, they represent firmly held convictions. Some of them may seem obvious, but they have been confirmed by our experimentation and we feel they are worth stressing. Others may seem counterintuitive or wrong. (In fact, some of our recommendations are in direct opposition to our preconceptions.) We despair of convincing the reader of any counterintuitive results through prose alone; we can only say, "You should sit down and use our experimental software systems." That is not a scientific or logical statement, but the design of man-machine interfaces and interactive systems unfortunately remains an art, not a science or an engineering discipline.

Issue #1: The Differences Between Paper Maps and Interactive Electronic Maps

There are obvious differences between paper maps and interactive computer-

based map display systems; indeed, this observation is one of the premises upon which our study has been based.

The most fundamental difference is the fact that the cartographer designing a paper map has only one shot at the user; he must put onto the map everything that the user—any user—will ever want from it. This single circumstance has broad ramifications. For example, it causes paper maps to be cluttered, so that the cartographer has the task of keeping the clutter within manageable bounds. Some of the key differences between paper and electronic maps are summarized below:

Item	Traditional Paper Maps	Interactive Electronic Maps
Cartography	Cartographic decision-making is done once by a professional cartographer.	The user acts as his own cartographer to some extent, but many constraints are removed and iteration toward a relevant display is possible.
Clutter	Tradeoffs must be made between clutter and the omission of needed information.	Information can be displayed selectively, so that vast amounts of data are potentially visible without clutter.
Communication	Communication via maps is limited to annotations and physical transmittal.	Electronic communication using map display as an I/O terminal is possible, as is effective narrow-bandwidth communication due to the rich context at both end-point sites.
Indexes	A printed index is common to allow the user to locate an object, given its name.	An electronic index can provide much richer information retrieval facilities; for example, it can identify and display the name of an object pointed at or it can emphasize all objects having specific attributes by blinking.
Physical parameters	Paper maps are very lightweight, portable, and inexpensive (in bulk quantities) and require no power source or communication link.	Size during next decade can approach approximately 10 cu. ft. (e.g., can be jeep-mountable); ruggedness is equivalent to that of telecommunication equipment; cost will be from $10K to $50K each (rough estimate).
Precision	Considerable precision is used within the map itself, since that is the only source of location data.	High-precision location information and other data attributes can be stored within the computer for access and computation, but high precision is not needed on the display screen itself.
Scale	Scale is fixed, with the possibility of insets of larger scale showing detail, or smaller scale showing overview.	Variable scale is possible, allowing access to potentially vast data collection on a single display surface; temporal insets (flipping in time between two different scales) are possible, as well as traditional spatial insets.
Symbology	Two-dimensional symbols are predominant; all symbols are static; color and size are used frequently to show data attributes; contour lines are commonly used to represent topography.	Three-dimensional and dynamic symbols are possible for representation of data attributes; blinking and change in intensity can be used to focus attention on data of interest; contour lines are unlikely to be the most effective representation of topography.
Text and legends	Text is heavily used to show names and data attributes, many fonts and type sizes are used to achieve clarity on a cluttered surface.	Text can be displayed selectively in response to requests and need not be continuously displayed; hence clutter is reduced and there is less need for elaborate fonts and character sizes.

This list is only an indication of *some* of the fundamental differences between these two map media. The main point we derive from it is the following:

GUIDELINE 1.1

Because of their fundamental differences, the design of electronic maps should not mimic that of paper maps. Each cartographic decision or design feature must be reconsidered, based on its underlying purpose.

Many of the features of electronic maps mentioned above are discussed in more detail below under specific issues.

Issue #2: Continuous Versus Discrete Display Controls

Imagine sitting at an interactive map display and being able to dynamically and continuously change the map scale from 1:1,000,000 (e.g., showing all of West Germany) to 1:100,000 (showing in some detail the town of Offenbach, West Germany) by twisting a knob. This provides some sense of the power of map display systems. In fact, however, we believe that in most situations this is a poor way to provide the user with what we have called "perspective control" over his display.

In the above scenario, the computer-based display system is being used as a camera; rotating a knob changes the focal length of a zoom lens. This is similar to the action of translation (panning a camera). Although it takes considerable computer power to recompute how the display should look each 1/30th of a second as the knob is being twisted, the computer is acting like a dumb machine, not a smart one; that is, its operation is not based on its potential knowledge of the data being displayed. In fact, this scenario is not even an efficient one. Consider what really happens when continuous control knobs are used for display control: After viewing West Germany, you decide to concentrate your attention on Offenbach. You turn the "zoom in" control, and soon Offenbach is disappearing off the upper right corner of the display screen (since the center of the zoom is, by default, the middle of the screen—in this case, a point somewhere near Frankfurt). So you begin turning the "translate up/down" and "translate left/right" knobs to center Offenbach on the screen. Then you zoom in again, and since your previous centering operation was only approximate, the city again begins to disappear along a display edge. By iterating this procedure several times, it is possible to get a close-in view of the region immediately surrounding Offenbach within about 10 seconds. But during those 10 seconds, the user was acting as a low-level control element in a closed-loop feedback system, and little thought was given to the real task at hand.

Now consider two alternative scenarios, both involving the computer's ability to manipulate and thus to some extent "understand" the symbolic information it is displaying, and the use of discrete, rather than continuous, display controls:

1. After viewing West Germany, you want to look at the area around the current field location of Headquarters, V Corps, which you know to be Offenbach. You type "Off" and are immediately presented with a list of data entities having that character string in their name. "Offenbach, W.Ger." appears, along with seven other West German cities. Using your tablet stylus, you move the display cursor to the particular entry on the list that you want, and that location on the map display begins blinking. You hit the "select" button (alerting the IMDS that Offenbach is an "active" data element) and the "go to" button (telling the IMDS to focus

the map display on the current active data element). The map display becomes centered on, and zoomed in on, Offenbach. (Note: it locates the desired point exactly, since the computer knows precisely where Offenbach is.) Upon viewing the display, you decide that you want a bit more surrounding context, so you hit the discrete "zoom out" button once, reducing the scale by a factor of about 1.3.[1]

You are now viewing Offenbach and its environs. The entire process has taken less than 2 seconds, and your attention has remained focused on the town, not on the process of getting there.

2. This scenario begins like the previous one, but since you are looking at the vicinity of Offenbach on the displayed map of West Germany, you place a displayed cursor controlled by your tablet stylus on the approximate location of the town, then hit the "go to" button, centering the display at your pen-indicated location. You then hit the digital "zoom in" button nine times in rapid succession, since you know from experience that three button pushes change the map scale by a factor of 2 (in this case, changing the original 1:1,000,000 scale to 1:125,000). Upon viewing the resulting display, you decide that you want a slightly larger scale, so you hit the "zoom in" button one more time, giving a final scale of about 1:100,000. This scenario takes about 3 seconds of elapsed time.

The second scenario points out an important design feature for digital display controls: The computer should be able to absorb and remember (in computer parlance, "stack") rapid successive button pushes, even if it cannot execute them as fast as they occur. In fact, the software should be designed in coordination with the interrupt mechanism, so that as soon as a new button push is received, the IMDS abandons work on creating the "old" map display image and starts displaying the "new" one that combines the effects of both previous function buttons. In this manner, the user need not wait for the IMDS to completely create all intermediate displays. Instead, he receives some partial intermediate feedback during his sequence of button pushes and gets the "final" resulting display soon after hitting the last button. This strategy has been used at Rand for a number of years in conjunction with a CRT-oriented text editing system called NED [8] and has been found to be very effective. We have therefore adopted and tested the technique on our map displays and again find it very effective and natural. The MITRE Corporation has also found discrete controls to be effective in its Geographic Data Display System [9].

The above discrete zoom scenarios generalize nicely: If the user enters a complex query using an underlying information retrieval system (e.g., "Where are all enemy SAM sites in sector 12?"), the IMDS automatically treats the data items flagged as a result of that query as being active; then when the user hits the "select" and "go to" keys, the display immediately changes scale and translates so that the geographic area shown on the display contains just those sites satisfying the request (those sites would also blink to focus attention on the currently active data elements).

The discrete zoom scenarios require considerably less computational power than the continuous zoom feature, since there is no requirement to maintain close

[1] Our experimental software systems have used a discrete scale factor of the cube root of 2 (= 1.25992105) for zooming in or out, since a factor of approximately that magnitude maintains user context and is not disconcerting, and it is easy to remember that three successive button pushes change the scale by a factor of 2.

real-time feedback to the user during a continuous process. In the event that the computer does fall behind in providing feedback to the user's actions, the process becomes awkward during continuous zoom; it is less aggravating if the computer is slow in responding to a discrete command.

We summarize the above findings as follows:

GUIDELINE 2.1

In a system with reasonable computational agents, continuous controls are in many situations considerably inferior to discrete controls such as function buttons.

There is an interesting phenomenon related to continuous controls that we feel is worthy of note. Which way do you turn a knob (clockwise or counterclockwise) to make the displayed map shift toward the right? It turns out that users have consistent preconceptions about the sense of direction of almost all continuous controls over a display. The interesting part is that the answer depends *consistently* on the size of the object being viewed in relation to the viewing surface. If the display is of the whole earth (e.g., as seen from outer space), then since the earth is small relative to the display screen, most users believe a clockwise turn of a knob should result in moving the displayed object (i.e., the earth) to the right. They feel that the knob moves the object, since small objects tend to be movable.

However, if the object is so large that only a portion of it fills the screen (as in most map displays), then there is a shift of user perception. The CRT screen becomes a window onto a large (and hence immovable) object, and the knob is felt to control the position of the window relative to the larger object. In this case, turning the knob clockwise is expected to move the window to the right, although the net result is that the map display moves to the left relative to the screen (since the display window—the CRT screen—is really fixed in space). Obviously, a discontinuity in the preconception is reached as the size of the object just fills the screen.

This effect may seem just an interesting bit of trivia, but users' preconceptions are quite consistent in this regard, and violations of those preconceptions are disconcerting. An Air Force captain using our system to solve a route-planning problem said, "Why is it that when I turn the knob up, the map goes down?" He firmly associated a particular rotation of the knob with the direction "up" and didn't even notice the dependence of his remark on that assumption.

We therefore conclude that:

GUIDELINE 2.2

Users have preconceptions that are quite uniform about the direction in which continuous controls should be moved to shift the display. Moreover, the expected direction changes—almost invariably—when the size of the object being viewed exceeds the size of the display window.

One major form of analog display control is what we have called "aircraft-type controls." By this, we mean allowing the user to "drive" or "fly" through a two- or three-dimensional space by the use of one or more joysticks or similar devices. More precisely, we use the term "aircraft-type controls" to mean control devices providing (1) continuous motion through space, (2) control over any derivative of position, such as velocity or acceleration, rather than position itself, and (3) a linkage between direction of motion and direction of sight. Such controls are natu-

ral for flight simulators; we wanted to understand their efficacy in allowing a user to change his viewpoint or perspective on an information display in other problem-solving and planning domains. Modern displays (such as the Evans and Sutherland Picture System 2) allow continuous real-time change of perspective as a function of user input. One of the potential advantages of such displays is that continuity of change in viewpoint provides context and orientation for the user during that change.

We had an excellent opportunity to investigate aircraft-type display controls through the generosity of Frank Lewandowski at Singer-Link Division, Sunnyvale. We were given access to an R&D version of a flight simulator that allows a user to "fly" within a computer-generated display of a significant portion of downtown Las Vegas. A typical scene generated by this device is shown in Color Plate 3. Subjects for our informal experimentation included a member of the Rand staff who has a pilot's license, a research assistant with less-than-average manual dexterity, a staff member with no pilot training but somewhat above average manual dexterity, and a staff member with below-average manual dexterity. The display control was a single joystick designed for flight-simulator experimentation, not for control of viewpoint in symbolic map displays; some of its characteristics were in fact quite inappropriate for viewpoint control. Nevertheless, we formed strong opinions about this type of display control that we feel have considerable generality. These opinions were reinforced by substantial experimentation with knob-type continuous controls for changing viewpoint (the type of control device on our in-house experimental equipment). Our primary conclusion regarding this type of display control is:

GUIDELINE 2.3

Aircraft-type controls are inappropriate for almost all geographic display applications.

This is a strong statement, but we believe it is justified. We of course exclude the flight simulators from the display applications referred to. As mentioned above, in the discussions of continuous versus discrete display controls, changing viewpoint with aircraft-type controls is more time-consuming than with other alternatives. Also, aircraft-type controls do not use the power of the computer to perform actions based on the content of the displayed information. We have come to believe that content-based directives are much more effective for changing viewpoint; for example, the user should be able to say, "Put me on top of *this* building, facing toward *this* point," rather than having to manually fly himself to that same position and orientation.

Three of the four subjects introduced to the rather complex joystick control of the Singer-Link system, including one subject with less-than-average dexterity, learned to maneuver adequately in three-dimensional space within 30 minutes. (The fourth subject was not capable of becoming an effective user. Mr. Lewandowski reports that this is true of about 5 percent of the people using his system.) Thus these controls would not be recommended for display applications involving untrained or sporadic users; however, intensive training is not required as a prerequisite to effective use. To summarize:

GUIDELINE 2.4

Aircraft-type controls can be learned in a few minutes by most subjects.

Issue #3: Clutter

The standard paper map is inevitably cluttered. As we have mentioned, the cartographer must present in a single display all the information that the user might need. Many cartographic tricks are used to control clutter, including multiple type fonts and character sizes, distinctive use of color, simplification, and insets.

The situation is quite different with interactive map display systems. It is theoretically possible to display vast quantities of information from a data base (e.g., the locations and names of all cities in the U.S. having a population greater than 5,000) with a single command to the system—thereby creating extraordinary clutter. However, it is also possible to give the user control over the amount of clutter through his interactions with the system. He can also display information selectively, so that the total amount on the screen remains manageable. The user is therefore in the highly desirable position of having much more data at his fingertips for potential viewing and computational use than he must ever see at once. We feel this is one of the most important characteristics of interactive map display systems, and one that should be capitalized upon in system designs. With today's technology, a "map" can contain literally millions of data items, all potentially available in almost infinite combinations yet manageable through the use of some simple clutter-control techniques.

Note that one natural way of giving the user control over the quantity of information displayed is through electronic overlays. (See items 3 through 7 under design Dimension #2 above.)

Because clutter control is such an important factor in electronic maps, we have addressed the question, Can the user control clutter effectively, or must such cartographic decisions be performed—albeit crudely—by the computer? We arrived at two relevant conclusions:

GUIDELINE 3.1

Given control over clutter, users tend to act responsibly and limit clutter effectively.

In one of our informal experiments, users were allowed to display a map of the continental United States, or any square portion of it. The display system automatically showed, as stylized abstract symbols, approximately 25 of the most important population centers and airports in the displayed region. (Figure 5 shows a representative display from this experiment.) In addition, the users could display—or remove from the display—any or all of the following:

- The continental boundary itself.
- All visible state boundaries.
- The names of all displayed cities.
- The names of all displayed airports.

We found that under these conditions, users did not overcrowd the screen with information. Even though they needed the information to perform a route-planning task, they acted very responsibly in controlling the amount of displayed information. Although our experience is limited to this single, controlled situation, the results were sufficiently consistent across the 12 users of the system for us to conclude that given simple, effective control over the amount and content of infor-

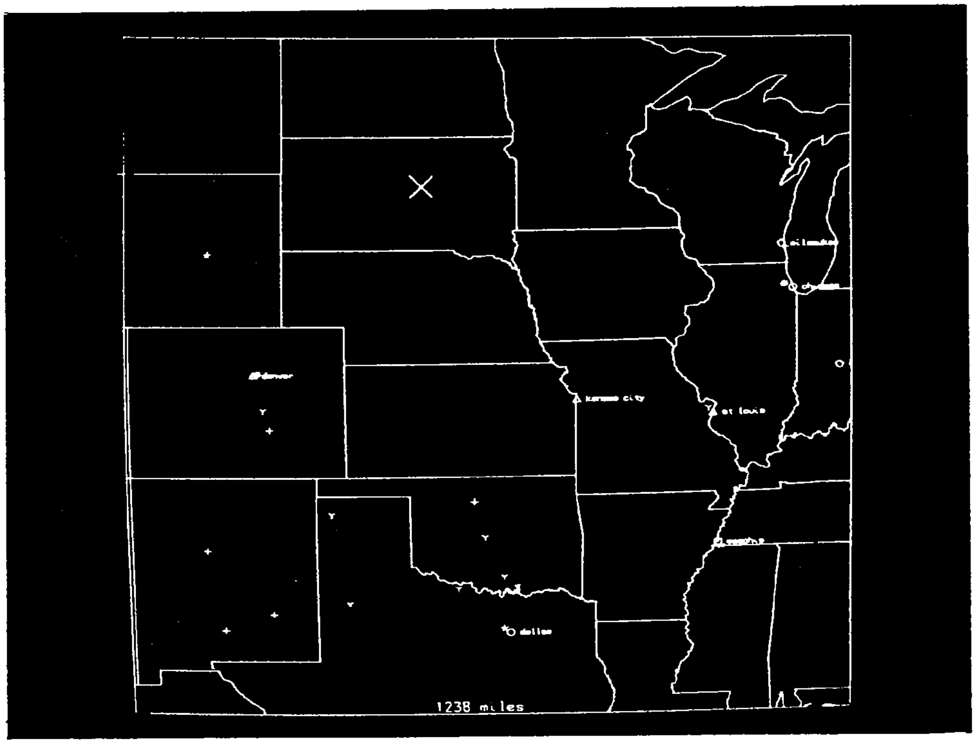


Fig. 5—Map display of a portion of the continental United States

mation displayed on an interactive electronic map, users will naturally tend to limit the amount of clutter on the screen. We therefore conclude that elaborate heuristic computer-generated clutter-control techniques are probably not necessary for most applications.

In this same experiment, we offered users the option of either seeing the names of all cities or airports (or both) displayed on the electronic map or pointing to a displayed data item and asking its name. In the latter mode, we simulated a voice-response unit by having an assistant read the desired information aloud from a CRT screen the user could not see. It should be noted that computer-generated voice-response output is a readily available technology, and it provides a parallel output channel for selective textual information that further eliminates clutter on the display screen. The results of our experimentation with approximately 12 subjects were consistent enough for us to state that:

GUIDELINE 3.2

Given a choice, users often prefer to receive voice-output data rather than CRT-presented data.

Issue #4: Orientation

Disorientation is not normally considered a problem in using traditional paper maps. A north-pointing arrow is usually conspicuous, and the user's orientation

with respect to the map does not change often or suddenly. Also, the amount of visible context is constant, since the scale does not change.

The situation is different with electronically displayed maps. If the user is permitted to zoom in, increasing the scale to the point where a single square kilometer (or meter!) of Europe is visible on the display, he might well forget which kilometer is being viewed (especially if he should be distracted). Similarly, the ability to rotate the displayed information or to "fly in" to a displayed region from an arbitrary direction can cause disorientation.

In our experimentation, we paid particular attention to the situations in which user disorientation occurred and to the frequency of this effect.

User disorientation did in fact occur, particularly during experiments with aircraft-type flight controls on the Singer-Link system. Probably the most common cause of disorientation was the lack of visible features on the screen. This occurred, for example, when the user was close to a (simulated) building and the entire display surface became filled with a single featureless wall of that building. (Disorientation also occurred when the display contained nothing but featureless sky or ground.) Until a feature appears, the user gets no feedback during rotation or translation of his viewpoint. An obvious solution is to use texture or features on any flat surface, but none of the displays we know of has texturing as a display primitive, so the provision of texture entails the use of considerable computation, memory, and other system resources. Designers of "movie maps" that contain computer-generated pictures of objects having flat featureless surfaces should be aware that user disorientation might well occur when user controls permit faces of these objects to fill the field of view.

GUIDELINE 4.1

Disorientation can be caused by abrupt changes in view, lack of visible features, and interruption of the user.

We found, however, that persons having a pilot's license exhibited considerably less disorientation from featureless displays than did nonpilots. We trace this to specific techniques taught as part of flight training (to prevent loss of orientation when encountering clouds, etc.); these techniques presumably can be taught to users as needed.

GUIDELINE 4.2

Disorientation can be reduced by specific training of users and occurs less frequently among trained pilots than among other subjects.

There is a set of display techniques that can be programmed to further reduce the risk of disorientation. For example, rotation of the displayed map might be disallowed, so that north is always up; or a north-pointing arrow might appear continuously in a corner of the display. The arrow could be shown in perspective to provide additional feedback if three-dimensional perspective is being used.

One technique we have found to be very effective is what we call a "recall" feature: Imagine that you are looking at the details of the topography between Hill 273 and Hill 138, and you become distracted and lose track of the context of that display—it's somewhere in the southeastern corner of West Germany, but what and where was the nearest town? If you hit the recall button, two things happen:

(1) a dashed box is drawn at the edges of your current screen display, and (2) the IMDS remembers your current display (i.e., its scale and location on the larger map) as being of interest. By hitting the "zoom out" button several times, you immediately see the context for your original display, with the area of interest outlined by a dashed box. As an example, Fig. 6 shows a display of the southwestern United States after the user had zoomed out from a detailed view of Southern

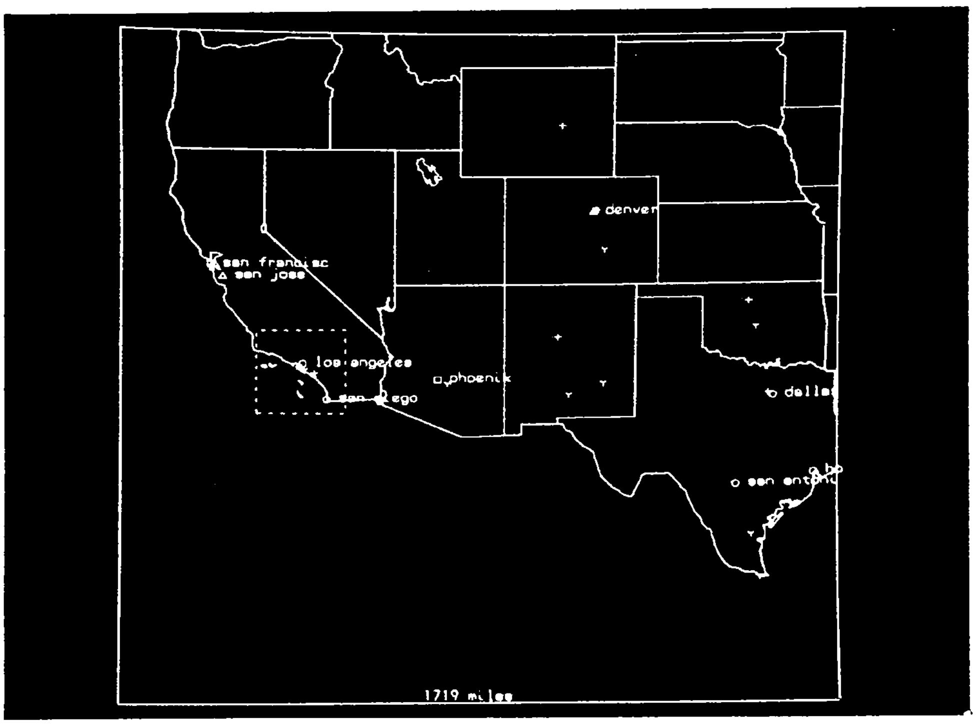


Fig. 6—Display of southwestern United States after zooming out from Southern California

California. Since the computer retains a knowledge of the area of particular interest, the user needs only to hit a single button to return immediately to his zoomed-in display—that is, to the area inside the dashed box filling the display screen. A simple extension of this technique allows the user to give a name to a particular displayed region, have a number of such named "reserved displays," and return directly to any of them by giving their name. (In a sense, this feature might be considered a "temporal inset" rather than a traditional map inset, since various user-defined full-screen insets can be accessed sequentially in time, rather than requiring portions of the display space to be used, as is normally the case in traditional maps.) Through such easily implemented features as these named insets, orientation can be maintained in a natural manner. Thus:

GUIDELINE 4.3

A variety of simple techniques can be used to retain user orientation in map displays; disorientation is not a significant problem.

Zooming or translating the display by a large amount can in itself cause disorientation. We find that our factor of 1.26 for zoom change (per button push) does not cause disorientation, but a scale change equivalent to five such simultaneous changes (equivalent to a scale change factor of 3.2) does, especially when zooming in. Similarly, a translation of 1/4 of the screen width does not cause disorientation, but an instantaneous, discrete translation step equal to the full width of the screen does. We conclude that:

GUIDELINE 4.4

Discrete zoom or translation increments greater than certain limited values cause disorientation and should be avoided.

Issue #5: Legends

A conspicuous feature of paper maps is the extensive use of embedded text to provide the names of most displayed objects. For example, consider the Automobile Club map of the west side of Los Angeles in Color Plate 2. It contains the names of all displayed surface streets, freeways, major points of interest, cities and districts, beaches, airports, mountain peaks, and other features—over 4300 characters in all! The layout of all that text, at varying orientations and with different character sizes and fonts, requires many subtle cartographic decisions. Is this much text necessary on electronic maps? If so, what are the implications for the system's character-generation capabilities and for the computation required to automate the cartographic layout problem?

We have found, happily, that interactive electronic maps do not need to mimic the complexity of paper maps:

GUIDELINE 5.1

Continuously displayed legends giving names of entities (and displayed text giving other properties of geographic entities) seem less valuable in interactive maps than in paper maps.

There are two primary text-related questions for which traditional maps provide answers: (1) What is the name of *this*?, and (2) Where is the entity named *x*? The first question is usually handled by text embedded in the map itself: The names of streets, cities, etc., are written near the entities. The second question is normally handled by a map index—an alphabetical listing of entity names (usually grouped by major category, such as roads, cities, sites of interest) with associated locations (usually grossly indicated by such designators as E5, meaning the entity is located near the intersection of vertical coordinate E and horizontal coordinate 5).

Not only can computer-based software systems within an IMDS almost always provide answers to both of the above questions much more efficiently than traditional paper map labels and indexes, they can also handle (at least) one other entire category of question that traditional maps cannot: Where are all the entities that meet condition *c*? The electronic equivalents of labels and map indexes that we have used to provide answers to these three major text-related questions are outlined below.

1. *What is the name of* this? This question is the simplest to handle. The user merely points, with a tablet stylus, for example, to a displayed entity; its name and

any related data attributes can appear on the screen, either at that location (which might contribute to clutter) or in a portion of the display screen reserved for textual responses and requests. (In our research, we used a separate alphanumeric CRT terminal located adjacent to the map display screen for such textual interactions.) In addition, it is useful to blink or intensify the object pointed to, providing feedback to the user regarding the object selected, its extent, and its spatial relationship to other nearby entities.

2. *Where is the entity named* x? To provide the answer to this type of question, we developed an experimental electronic map index program that we feel is considerably superior to traditional printed map indexes. As an illustration of its operation, consider the following scenario:

Col. Jones has received a garbled intelligence report stating that an enemy command post has just been established at the northwest corner of the village of (what sounded like) "Shinefurt." The coordinates were garbled, but it is somewhere in Sector 12, which contains approximately 1200 small towns and villages. From his familiarity with the German language, Col. Jones believes the true village name could be either Schweinfurt, Schweinfort, Sheinfurt, Sheinfort, or possibly Kleinfurt, Reinfurt, or Reinfort. With a traditional map, he would look up each name in an alphabetic index, then search the subsector given for each place, then ascertain from other conditions (e.g., whether that territory is in enemy or friendly possession) if that candidate is a possibility, then decide which of the remaining candidates is the most likely one. With an electronic map index, he simply types a specification to the computer, such as

(S K R)*einf(u o)rt

meaning he would like to see all place names starting with either S, K, or R, followed by any (including zero) letters, followed by the character string "einf" followed by either u or o, followed by rt. All place names meeting those conditions immediately begin blinking, and by hitting two function buttons (one to select those entities as being of interest, and one to zoom in on the selected entities) he brings into his display screen the map region showing all candidates meeting his conditions.

Our experimental program included another helpful feature, as well: As each letter or clause of the user's specification is typed, all place names meeting that partial specification are displayed on an alphanumeric CRT terminal. With a 9600 bit per second connection, the display is sufficiently rapid that this feedback does not interfere with normal typing. Therefore, as the specification becomes formulated, the list of candidates meeting the partial specification decreases until it can be scanned by eye; the user can then stop typing and either select the entire remaining list (by hitting a carriage return) or move the cursor to the particular entry he wants and select it.

In our tests, using lists of all the street names in Santa Monica and Bel Air, California, the user could find the location of a street whose name was imperfectly understood and have it centered on the displayed map and blinking within 3 or 4 seconds. An equivalent process using a printed map index takes at least one, and often two, orders of magnitude more time and imparts less information to the user. (We can illustrate this last point by the following example: The extent of streets is often not clear from printed maps, but since the computer data base contains both

the extent of a displayed object and its associated name, the total object having that name can be blinked, making its extent and its relationship to surrounding objects immediately clear.)

3. *Where are all the entities that meet condition* c? This is really an extension of question 2. We assume that the data base supporting a map display system contains named objects, each having associated named attributes and corresponding values. For example, it might contain the following two objects:

Population Center		Bridge	
Type:	village	Construction:	steel-beam
Name:	Schweinfurt	Capacity:	20 ton
Population:	1270	Lanes:	2
Lat:	478525	Type:	highway
Long:	038360	Lat:	477330
		Long:	039125

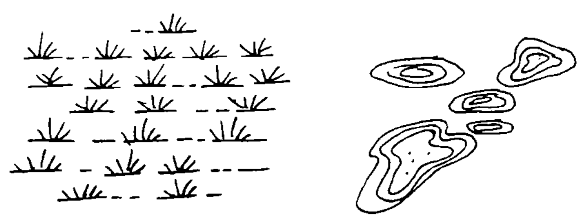
With these data, finding all objects satisfying a stated set of conditions—for example, all highway-type bridges in sector 12 that have capacity greater than 15 tons—becomes a straightforward information retrieval process. (We assume that the information retrieval program has access to auxiliary information giving the coordinates of the boundaries of sector 12.) Again, all objects (i.e., bridges) meeting this condition can be blinked, and the IMDS can itself zoom and scroll the display so that all of them are in the field of view. We conclude from these experiments and their natural extensions that:

GUIDELINE 5.2

Electronic map index programs can greatly increase a user's ability to locate information on a map. They depend, however, on a system design in which the computer acts not just as a camera, but understands the names and attributes of the data being displayed.

Issue #6: Variable Abstractions

Different symbols are used to represent identical information on traditional maps, depending on the scale of the map. For example, two common symbols for marshland are:



Neither symbol is appropriate at all scales, and neither can be replaced at all scales by an aerial photograph of the marshland being viewed. Also, the abstractions at various scales do not merge gracefully into a continuum; there are discontinuities.

There is another important way in which abstractions change as a function of scale: through the process of selection. When a road network is depicted on a

1:500,000 electronic display map, only major routes are shown, and they are quite stylized (e.g., widths are exaggerated). As the user zooms in, it becomes appropriate to begin showing certain secondary roads, then more of the secondary road network, and so forth. If the zoom is continuous, then it becomes necessary for symbols to appear (and in the reverse zoom, to disappear). This causes additional discontinuities in the display.

The question becomes, How disconcerting are these changes in the abstractions to a person attempting to use an electronic map for planning or problem-solving? We developed an experimental map of the United States containing over 4,000 point entities (populated areas and airports), with capabilities for continuous zoom and scroll control by the user. Only about 25 such entities were displayed at any one time, but the symbols representing them appeared, disappeared, and changed form as the user changed scale or scrolled the display. (Changes in the form of the symbols indicated *relative* importance of the entity on the current display, and this relative importance changed, even during scrolling at a constant scale.)

Twelve users performed a complex route-planning task with this display in multiple sessions. The results were consistent enough for us to state the following with some assurance:

GUIDELINE 6.1

Users can tolerate and effectively use variable abstractions, provided they can control the abstraction process.

That is, the user should have enough control over the abstraction process, and over controlling parameters such as scale and translation, that he can regulate the process giving rise to the changes in abstraction. Another consistent result was that:

GUIDELINE 6.2

Users' abilities to tolerate variable abstractions increase with experience.

Some users' first reaction to a map with symbols that appeared, disappeared, and changed form in response to their actions was one of horror. However, essentially all users acclimated rapidly and began using the information contained in these changing symbols effectively. Changing abstractions present less of a problem with discrete zoom control than with continuous control (see Issue #2 above), since abstraction changes can occur at natural discontinuities in the display.

Issue #7: Electronic Maps as Problem-Solving Aids

As we previously indicated, electronic maps should really be thought of as aids to problem-solving or planning activities. Their true power comes from their role as an interface to an assortment of computational tools, in which both arguments to processing functions and results of those functions are two-dimensional and graphical.

The following examples illustrate this philosophy:

- "What is the best route from *here* to *here* capable of handling a 15-ton tank and avoiding *this* village?"

This request calls a network route-optimization program with relevant parameters and constraints; upon receipt of the solution, the display intensifies or blinks the selected route to highlight the information.

- "If we fly at 800 feet from *here* to *here*, then climb to 2500 feet and proceed to *here*, where should the refueling rendezvous be if the tanker takes off from *here* at 0830 hours?"

  Again, this request provides an effective interface for information interchange with a complex optimization program.

- "Show me the locations of all targets hit during the period 25 June 1978 –25 October 1978."

  This is an information retrieval request whose answer is best displayed graphically to provide the user with a gestalt of the battle action (which could involve thousands of data items) over the past three months (see the discussion of an IMDS as a window onto a data base retrieval system under question 3 of Issue #5 above).

Such interactions need not require the system to interpret English-language statements. A much more stylized language, such as:

"route: begin <here>, end <here>,
constraint = avoid <here>,
constraint = capacity .ge. 15t;"

can be used, especially if the system is used by trained professionals who use it often and intensively. We have used English-language queries merely so that their intent is understood by readers with diverse backgrounds and interests.

In our experimentation, we accessed both information retrieval and computational facilities via an electronic map. We used such facilities intensively to clean up inconsistent and incomplete geographic data files which were provided by various sources. A strong conclusion emerging from these activities is that:

GUIDELINE 7.1

Interactive maps are effective problem-solving devices, even when supplied with rudimentary computational and information retrieval facilities.

IMPLICATIONS FOR SYSTEM ARCHITECTURE

These guidelines provide a number of implications regarding the architecture of an interactive map display system. We have not studied system design issues in depth, so our discussion must be treated—like most of this report—as a set of indications worthy of further consideration, not as fact or dogma. With this caveat, we present the following inferences for designers of map display systems.

Data Base Design

Interactive map display systems require potentially vast amounts of data in order to be effective. For example, consider an electronic map of West Germany that permits the user to zoom into any square kilometer of the country and view

50 separate data items (such as villages, bridges, or roads), with each data item requiring twenty-five 32-bit computer words of information to store its name and associated attribute values. Assume that the relevant topographic information regarding the country is stored as a 50-meter-interval grid, with each gridpoint represented by three computer words of information. The data base for this map would therefore require a total of more than 6×10^8 (six hundred million) words of computer storage,[2] or about 2×10^{10} bits.

Various data-compression schemes might be used to reduce this total by an order of magnitude or so. However, it is clear that the organization of geographic data bases is a crucial design issue for interactive map display systems. (We assume for this discussion that all map data are stored in computer-manipulable form, rather than on back-projected slides or overlays, for the reason given earlier: The system is much more powerful if it knows the attributes of all the data it is displaying.)

If the user has complete freedom to zoom or scroll his display within a wide geographic area, including the ability to increase the scale until his display shows 1 square kilometer or less, then the demands on the organization of the data base are severe—most likely they are unattainable with present technology—since any portion of the data might quickly be called into view by the twist of a knob.

There is some serendipity in our recommendation that discrete zoom and translation controls be used (Guideline 2.1). In a discrete-control system, data can be stored in geographic "pages," with the page currently within view and its eight nearest neighbors brought into high-speed memory. In any single zoom or scroll button push, those nine pages of data should suffice to create the required display. If the field of view changes to the point where new neighboring pages might become visible, they can be loaded into high-speed memory in a buffering scheme that operates in parallel with responses to user actions. It is also possible to precompute and prestore selected map abstractions at various scales (also in geographic page form), with the digital zoom controls arranged so that these preselected scales are the only ones accessible by the user. For example, such precomputed abstractions might be stored for scales of 1:100,000; 1:200,000; 1:400,000; and 1:800,000. If each zoom button push caused a scale change of a factor of about 1.26 (actually, the cube root of 2), then after each three successive button pushes, a new abstraction level would be used in computing the current display, and only the geographic data pages for that abstraction level need be accessed. Hence, although a voluminous data base is still needed, digital display controls put constraints on the user's freedom to access those data and allow the use of paging schemes to limit the amount of data in active use at any one time. (A recent implementation of a data paging scheme for geographic display systems is described in Lehman [9].)

We emphasize, however, that our preference for discrete display controls does *not* derive from their relative computational efficiency; we would recommend their use even if it entailed a considerable computational overhead.

[2] (50 data items/sq km) $\times$ (248,417 sq km) $\times$ (25 words/data item) = 3.11×10^8 words; (400 gridpoints/sq km) $\times$ (248,417 sq km) $\times$ (3 words/gridpoint) = 2.98×10^8 words; $(3.11 + 2.98) \times 10^8 = 6.09 \times 10^8$ total words of storage required.

Data Tablets

Throughout this report, we have depicted the user's commands to the IMDS as being of the form:

"Put the center of the display "<here>."
"What is the best route from <here> to <here>?"
"What is the name of <this> bridge?"

In our research, we have been using a data tablet and stylus to move a cursor on the display screen and to record location information (e.g., *here)* when the pen is depressed on the tablet. (Depressing the pen closes a microswitch in the tip and sends an interrupt to the computer.) Although a trackball or joystick might be used to perform similar actions, we have found that the flexibility of a tablet and the possibility it affords for drawing and editing routes make it very attractive for IMDS purposes. On the basis of our experience, we recommend that a data tablet be considered part of a standard IMDS design. As an alternative, a touch-sensitive display surface might suffice for many applications.

Display Precision

Consider a user performing a route-planning task on a 1:1,000,000 scale map of West Germany. Since 1 inch on the map equals 1,000,000 inches (15.8 miles) of terrain, a measurement error of only 1/8 inch on the map translates into an error of about 2 miles. Because of the large multiplication factor for measurement errors, maps are drawn with considerable precision.[3]

The situation is completely different with electronic maps, since scale change is always available; it is in fact possible to design systems in which a switch from small scale (for gross planning) to large scale (for precision in locating endpoints of route segments, target locations, etc.) is available at the touch of a button. Rather than be accurate to 1/8 inch on a 1:1,000,000 map, the user can be accurate to 1/8 inch on the same map data at a scale of 1:10,000 (giving an error of less than .02 miles, or about 104 feet). Furthermore, since it is assumed that the computer knows about the underlying data, the user can point imprecisely to a known location (such as a crossroad or target site) but have the computer make its computations using not his approximate pointing location, but the exact location of that site to the precision stored in the data base.

We therefore believe that in most applications of electronic maps for command and control planning and problem-solving activities, normal commercial TV display resolution (i.e., approximately 512×512 picture elements) is sufficient. As resolutions such as 1024×1024 become available at modest cost, they will most likely become the standard for IMDS designs, not so much because the precision is needed, but because the resulting picture is aesthetically more pleasing (for example, there is much less of a "staircase effect" in displayed diagonal lines).

[3] Since they could be constructed with even more precision than is customary, it is interesting to consider how the amount of precision is chosen; we speculate that the answer involves the length of the human arm (affecting the overall size of a typical map), the distance at which the human eye can comfortably focus (affecting the minimum useful field of view), and the resolution of the eye (affecting the amount of data that can be usefully displayed within that field of view).

Bit-Map Versus Calligraphic Displays

We have experimented with both black-and-white calligraphic (i.e., line-drawing) displays and with bit-map color raster scan displays. We have developed no strong preference and believe the choice depends heavily on the application. If coverage area swaths for sensors and satellites must be depicted, bit-map displays are probably preferable, because irregular areas can be "colored in"; however, current bit-map displays require seconds of elapsed time to change the $512 \times 512 = 262{,}144$ pixels of display, whereas the calligraphic display can respond in a time that appears instantaneous to the user. Bit-map color displays will most likely be the ultimate choice for IMDS displays, but the choice will not be an obvious one until the entire screen contents can be updated in about 1/10 second or less.

Rear Projection of Slides, Static Overlays, and Videodisks

Several current map display systems being tested and evaluated for command and control applications display computer-generated data overlaid on map images obtained through such techniques as rear-projection of photographic slides. In the not-too-distant future, it will be possible to store digital or analog map images on videodisks and use them in a similar manner.

The resulting displays will retain much of the appearance of traditional paper maps but will allow dynamic presentation of current data. They therefore will present little "culture shock" to users familiar with traditional maps. Through judicious use of computer-controlled zoom lenses or equivalent digital techniques, it will even be possible to give the user some control over scale and translation on these map displays.

As the discussion of Guideline 2.1 indicates, however, considerable power is obtained in map displays through having the computer cognizant of all data displayed. When the user wants to zoom in on *this* bridge, or target *this* crossroads, or find out the population of *this* town, the power of the IMDS comes from having the computer know which data item, not just which approximate coordinate location, is being accessed. We therefore believe that slide- or videodisk-based map display systems are a useful, but only transitional, step toward the type of IMDS we have been discussing in this report. They will fill a need until the large and detailed geographic data bases required by a data-based IMDS become available.

Communication Links and Map Display Systems

Having a single, isolated map display system is like having a single, isolated intelligent terminal. It is useful, but it is not achieving its true potential.

For example, map display systems sharing a communication link can provide an excellent low-bandwidth communication medium between sites. As mentioned earlier, a company commander might link his IMDS to his battalion headquarters and create the following message:

> "Present location <here>; expected route of advance tomorrow to <here> by 0930, then via <here> to <here> by 1315 hrs."

The resulting digital message transmitted from company headquarters to the battalion headquarters can contain the text of the above message, with coordinate pairs substituted in the appropriate places. Assuming the coordinate information is coded so that it is easily read by an IMDS receiving this message, the battalion IMDS can display the message and can (for example, upon being pointed to by the battalion user) display the planned route of advance graphically on the battalion-level map. Note that it is not necessary that the company and battalion maps be displayed at the same scale or show all the same information. Each system can turn screen coordinate information (e.g., *here)* into a standard military coordinate system, and vice versa. Although the route of advance might have filled the company IMDS display screen, it might be only a portion of a larger display at the battalion level, where all the routes of advances can be studied as a single pattern. Note also that there is less chance of error in having a user point to a location on a map than in having him type or verbally dictate from 12 to 16 coordinate digits to transmit the same information.

IMDSs can also provide a medium for group coordination of plans, as well as a variety of other communication tasks. We conclude that the IMDS designer should take into consideration its possible roles as a type of data terminal in a communication network. The protocols for data transmission among IMDSs, and among an IMDS and other remote information systems, should be given more consideration than they have, to our knowledge, received to date.

III. CONCLUDING REMARKS

In a sense, the conclusions we have reached regarding the potential uses and capabilities of interactive map display systems are scattered throughout this report; they are stated in Sec. II as design guidelines and as points of interest along five different design dimensions. Our primary conclusions, however, can be summarized briefly as follows:

- Interactive map display systems can be extremely effective planning and problem-solving aids.
- Their design should not mimic traditional paper maps; they are a completely different medium whose design must evolve from their role as aids to the fundamental problem-solving activities of the user.
- The role of abstraction and the means given to a user for controlling it are both essential in the design of interactive map display systems.
- Most of the power of computer-based map display systems derives from having all geographic data in a form that allows computer interpretation of those data.

Because of the informal nature of our experimentation, these guidelines and conclusions must be subject to further evaluation. However, we have demonstrated some dramatic increases in problem-solving effectiveness using only rudimentary software systems. Further development, test, and evaluation of such systems, tailored to particular real-world problem-solving activities, seems warranted.

Appendix

MAJOR INTERACTIVE MAP DISPLAY ENVIRONMENTS USED IN RESEARCH EXPERIMENTATION

In our experimentation, we used three major map display environments to test and evaluate a number of hypotheses concerning interactive map display systems. We use the term "environment" to indicate a combination of software, data bases, hardware, user interface peripherals, and scenarios that together form an interesting portion of a complete map display system. As mentioned in the body of this report, none of these environments formed a complete planning or decision-aiding tool for a realistic problem domain. Each was designed to test specific features and hypotheses, sometimes by quite unnaturally stretching the limits of that particular design feature.

This Appendix provides a more complete description of each of these experimental environments, as additional context for evaluating the design dimensions and guidelines we have presented. These descriptions are of necessity brief and sketchy; readers wishing more detailed description should contact one of the authors.

1. ROUTE PLANNING ON U.S. MAP OF CITIES AND AIRPORTS WITH VARIABLE ABSTRACTION

The user was initially provided with a displayed map of the continental boundaries of the United States, along with symbols representing the location and importance of the n most important cities and airports. (Importance was determined by population for cities and by a function of runway length for airports.) In our experiments, the parameter n was usually 25.

The display apparatus was an Evans and Sutherland Picture System 2 that provides high-resolution black and white line drawings and symbols on a 25-inch calligraphic (i.e., directed-beam as opposed to raster scan) CRT. An E&S dial box containing eight continuous rotation dials was available; three of the dials could be used to control zoom in-out, and display translation up-down and right-left. A Summagraphics data tablet with stylus was available for route drawing and pointing; a displayed cursor indicated the position of the stylus. In addition, an Ann Arbor alphanumeric display terminal with keyboard was located adjacent to the E&S display screen, for display of auxiliary textual information. In most experiments, the Ann Arbor terminal was used by a research assistant to control portions of the display upon verbal command by the user, and portions of the information displayed upon the Ann Arbor screen were spoken by the assistant to the user upon request, to simulate a computer-driven voice response unit. The keyboard of the Ann Arbor terminal could be used as a set of function buttons to regulate the display contents. The primary function buttons available were:

Z-IN: Zoom in by a factor of 1.26 (see footnote, p. 22).

Z-OUT: Zoom out by a factor of 1.26.

Z-IN5: Macro zoom in—equivalent to 5 successive Z-IN button pushes (i.e., a total zoom factor of 3.17).

Z-OUT5: Macro zoom out—equivalent to 5 successive Z-OUT button pushes.

RECALL: Draw a dashed box around the current screen border and remember this display for later recall.

RETURN: Return to the display indicated by last "recall."

OVER: Return to the original (overview) display of the total continental United States.

SELECT: Select an <x,y> point by tablet stylus.

CENTER: Center the map display on a selected point.

LEFT: Translate left by 1/4 of the screen width.

RIGHT: Translate right by 1/4 of the screen width.

UP: Translate up by 1/4 of the screen width.

DOWN: Translate down by 1/4 of the screen width.

CNAME: Display the names of all visible cities (on/off).

ANAME: Display FAA 3-letter designations of all visible airports (on/off).

SBOUND: Display all visible state boundaries (on/off).

CBOUND: Display continental boundary (on/off).

DRAW: Draw a route by displaying line segments connecting successive locations indicated with the stylus.

SEGSEL: Select a segment of a drawn route with the tablet stylus for editing.

SEGDEL: Delete a selected route segment.

SEGEDT: Edit a selected route segment by respecifying one or more midpoints.

QUIT: Return from the DRAW or route segment editing mode.

WHAT: Select a displayed city or airport symbol by pointing, for use in an information retrieval request regarding that object.

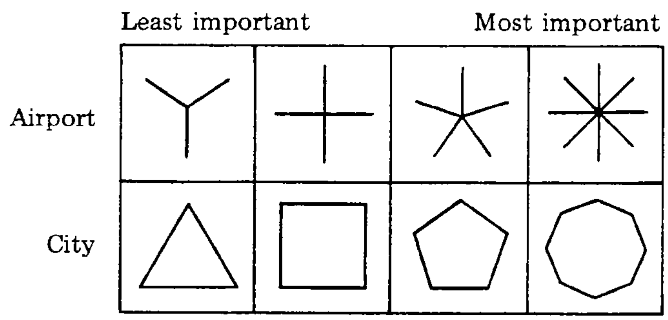
LENGTH: Select two points on the display, and the distance between them (in miles) will be shown on the text display.

CLEAR: Clear the screen of all drawn routes and dashed boxes.

The discrete function button zoom and translate controls and the continuous knob controls were simultaneously operable, so that either could be used, interchangeably, to affect the display. The current map scale was indicated at all times by a legend at the bottom center of the display giving the distance (in miles) across the visible screen.

Approximately 4,000 of the most important U.S. city and airport locations comprised the data base for this map environment. In addition to the coordinate location of each city and airport, our data base contained the name of the object

and a measure of its importance. As the user changed the geographic area being displayed through zoom or translation, the n most important objects for that new geographic area were computed, causing some cities or airports to suddenly appear on or disappear from the display. The abstract symbol used to represent an airport or city on the display was chosen from a family of four similar symbols to reflect the index of importance of each object relative to other objects currently on the screen. We call this use of relative symbology "variable abstraction." One of the symbols sets used is shown below:



A primary goal of this map environment was to test the effectiveness of variable abstraction in conveying information to the user (and, conversely, to test the degree to which such a display feature might be distracting or annoying).

We conducted an informal experiment in which each of 12 subjects was given the following task:

> You are to draw a route to be flown within the continental United States by a new experimental aircraft. You are to visit, or come as close as possible to, all major airports. The more important the airport, the closer your route should come. This route should also avoid all major population centers; the more important the population center, the more critical is your avoidance. The total route length cannot exceed 12,000 miles. Do the best job you can. There is no time limit.

This task was designed to require use of both overview and detailed map displays, so that the variable abstractions would be manifested during these operations. Each subject performed this complete task in each of three or four separate sessions and was thoroughly debriefed after each session regarding his attitude toward and assessment of various display features.

The authors also experimented extensively with this display environment.

2. URBAN ROAD NETWORKS AND ELECTRONIC MAP INDEXES

For the second display environment, machine-readable representations of the street networks for the cities of Santa Monica and Bel Air, California, were extracted from Bureau of the Census DIME files. Individual streets were displayed as a sequence of line segments, as shown in Fig. A.1.

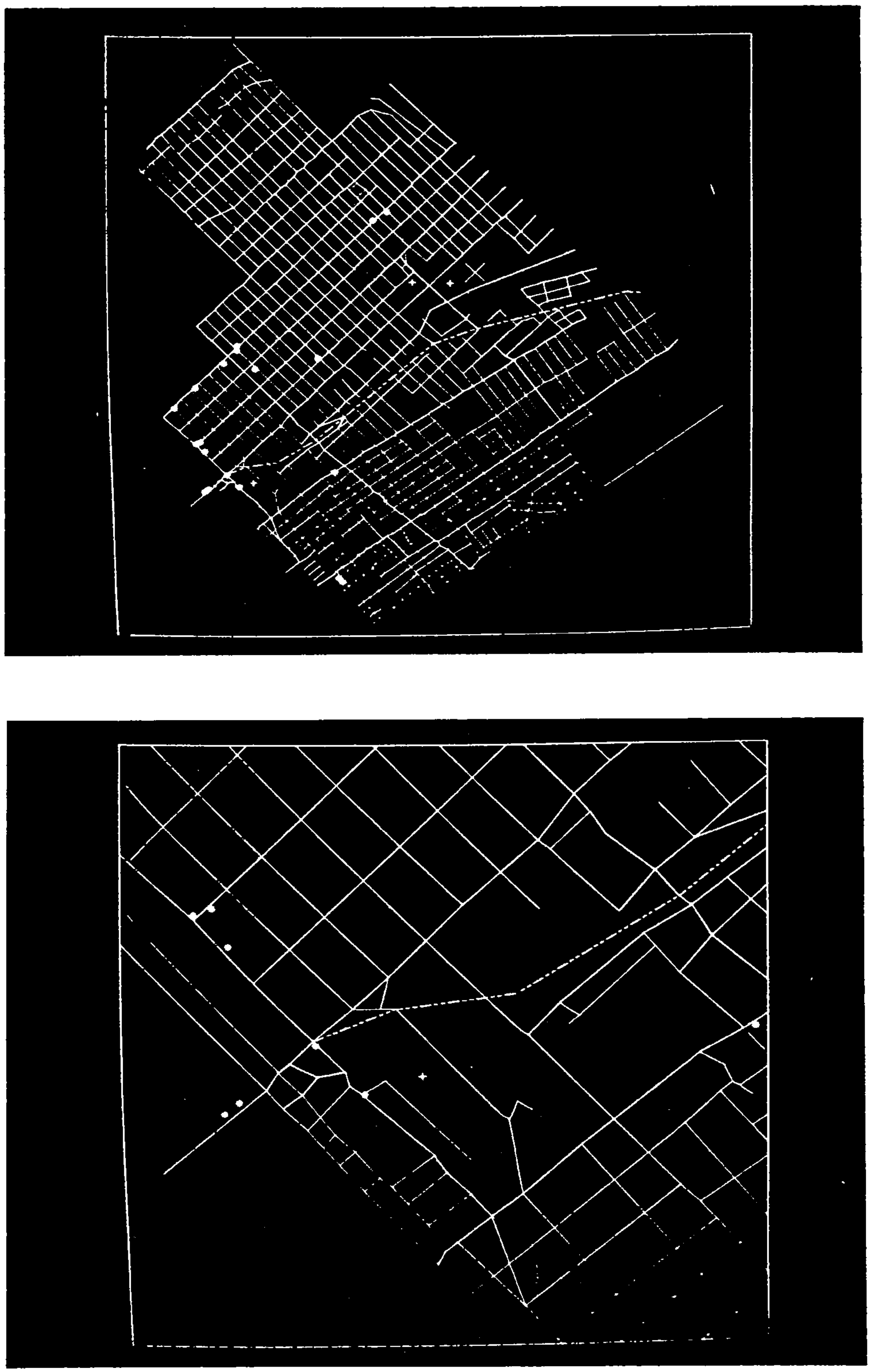


Fig. A-1—Displays of Santa Monica street network

The display apparatus for this environment was the same as for the first environment, described above, with the exception that a display could be copied from the E&S calligraphic black and white display to a bit-map color raster scan display system (manufactured by Genisco, Inc., of Irvine, California). The resolution of the Genisco display was 512 $\times$ 512 picture elements. We could therefore compare a high-resolution calligraphic display with a moderate-resolution color bit-map display of the same information.

All the discrete function buttons of the first environment were again available, except those having specialized application to the map of the continental United States (e.g., those controlling display of city and airport names, and state and continental boundaries). Control knobs and tablet were available as before.

A primary goal of this environment was to test the need for continuously displayed names of objects (such as streets) when other alternatives for accessing name information were available. To test the extreme case, the system did not show any textual information on the street network display. All street name information was accessed via electronic map index programs specially designed for this environment. The Ann Arbor text display terminal was used to display the name of any street pointed at with the tablet stylus (at the same time, the selected street began blinking on the CRT display to indicate its extent). The text terminal was also used in a highly interactive mode to show the names of all streets satisfying a pattern specification, as discussed on pp. 29-31, following Guideline 5.1.

One additional function button was programmed, tested extensively, and found to be very effective:

> Z-SELECT: Zoom in on the data items selected by the tablet stylus or by the electronic index feature.

Selected items were blinked on the CRT screen to highlight them, and it was found very convenient and natural to direct the computer to zoom and translate the display automatically to a point where exactly those selected items filled the screen.

This map display environment was used intensively by one of the authors (Shapiro) to perform extensive editing on the original DIME file information. (The files contained 1970 data with numerous errors; they were corrected and updated to reflect the current street network and to include points of interest.)

3. REAL-TIME INTERACTIVE FLIGHT SIMULATOR WITH HIGH-RESOLUTION COLOR AND HIDDEN SURFACE REMOVAL

Through the generosity of Mr. Frank Lewandowski of Singer-Link Division, Sunnyvale, California, we were able to conduct a series of informal experiments on an advanced research flight simulator developed by Singer-Link. Some of the key features of this system are:

1. *Very high-quality resolution and optics.* The system has an RCA 1024 $\times$ 1024 bit-map color CRT display, with reflection optics using a parabolic mirror to present a large display surface (about 30 $\times$ 40 inches) having the illusion of depth.
2. *Very high-speed display.* The entire display is updated each 1/60th of a

second, providing complete real-time feedback to the user's display control actions.

3. *Simulated three-dimensional perspective picture.* The display has colored surfaces, hidden surface elimination, and many other advanced features (but does not include computation of shadows or variable reflectance). Color Plate 3 is a representative picture generated by this system. The picture shows a portion of downtown Las Vegas.
4. *No abstractions.* The display produced by this system is very realistic, with little abstraction.
5. *Aircraft-type controls.* Devices provided for user control over the display comprise a single joystick that controls velocity and direction of motion (with direction of view linked to direction of motion). An auxiliary speed control knob provides a "multiplier" for the velocity control on the joystick. The software system simulates airplane flight parameters such as momentum and minimum turn radius.

REFERENCES

1. Lawrence, G.R.P. (ed.), *Cartographic Methods*, Methuen & Co., 1971.
2. Raisz, Erwin (ed.), *Principles of Cartography*, McGraw-Hill Book Co., Inc., New York, 1962.
3. Robinson, Arthur H., and Barbara Petchenik (eds.), *The Nature of Maps*, University of Chicago Press, Chicago, 1976.
4. Topfer, Friedrich (ed.), *Kartographische Generalisierung*, Geographisch-Kartographische Anstalt Gotha/Leipzig, 1974.
5. Davis, John C., and Michael J. McCullagh (eds.), *Display and Analysis of Spatial Data*, John Wiley & Sons, Inc., New York, 1975.
6. Safran, Navad, "Trial by Ordeal: The Yom Kippur War, October 1977," *International Security*, Fall 1977.
7. Sacerdoti, Earl D., "Language Access to Distributed Data with Error Recovery," in *Proc. 5th International Joint Conference on Artificial Intelligence—1977*, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1977, pp. 196-202.
8. Bilofsky, Walter, *The CRT Text Editor NED—Introduction and Reference Manual*, The Rand Corporation, R-2176-ARPA, December 1977.
9. Lehman, D. H., *Geographic Data Display Implementation*, MITRE Corporation, AD-A044 621, June 1977.

RAND/R-2382-ARPA